CHARACTERIZING THE EFFECTS OF CHRONIC CONVENTIONAL HEAVY CRUDE OIL EXPOSURE ON THE GROWTH, DEVELOPMENT, AND BEHAVIOUR OF LARVAL WOOD FROG (*LITHOBATES SYLVATICUS*)

by

Adam L. Scott

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Environment and Geography

University of Manitoba

Winnipeg, Manitoba

ABSTRACT

In recent years, the need for data on the effects of oil spills on freshwater environments has become increasingly important in Canada. This thesis will address some of these gaps by focusing on the ecological impact of oil spills on amphibians, specifically wood frog (*Lithobates sylvaticus*) tadpoles, within the context of two major research studies: the Freshwater Oil Spill Remediation Study (FOReSt) and the Floating Wetland Treatments to Enhance Remediation (FloWTER) study which took place at the International Institute for Sustainable Development Experimental Lakes Area (IISD-ELA) in Northwestern Ontario. The FOReSt study aimed to assess the effectiveness of non-invasive oil spill remediation techniques in shoreline habitats, while the FloWTER study investigated the potential use of symbiotic plant-microbe systems via engineered floating wetlands for the degradation of oil-derived hydrocarbons. FOReSt simulated an oil spill event and industry-standard remediation response (shoreline rinsing), characterizing the fate and bioaccumulation of oil components in tadpoles exposed to weathered conventional heavy crude oil (CHV) in shoreline enclosures. Results indicated significant bioaccumulation of total polycyclic aromatic compounds (tPACs) in tadpole tissue, though no measurable effects on growth, development, or survival were observed. Qualitative histological analysis of hepatic tissues found no significant differences between treatment and control groups, suggesting that tadpoles did not experience adverse health impacts under the conditions of the study. The amphibian component within the mesocosm tanks explored the chronic exposure of wood frog tadpoles to a water-accommodated fraction of CHV in combination with floating wetland treatments. Behavioural assays and histological evaluations were conducted to determine the impact of these oil spill response techniques on tadpole health and behaviour. The results showed no significant differences in bioaccumulation, growth, development, or behavioural metrics, such as total distance travelled, or time spent near conspecifics. These findings suggest that the amphibian early life stages studied were resilient to the oil spill scenarios tested under the conditions and remediation methods applied. Further research is needed to fully understand the long-term impacts of oil spills on amphibian populations and freshwater ecosystems.

ACKNOWLEDGEMENTS

To my advisors, Dr. Mark Hanson and Dr. Jose Luis Rodriguez Gil, your unwavering support and guidance have been the cornerstone of this journey. Your confidence in me and this project has been a beacon of strength. Together, we braved a global pandemic, a university strike, a field season forest fire, and an across-Canada move. Through it all, I gained invaluable skills and life lessons under your mentorship.

To the University of Manitoba, the project funders, my Master of Science committee members, and the Stress Ecology Lab members—thank you for your steadfast support and collaboration. To the municipality of Tub Town and the incredible staff at IISD-ELA, thank you for your guidance during fieldwork. Your commitment was critical to this project's success. A special thanks to the wood frog tadpoles (and toads), whose cooperation and ultimate sacrifice made this research possible.

A heartfelt thanks to the 2021 Toxicology Team (Grubs): Aidan 'Aido' Guttormson, Blake 'from chem lab' Cooney, Carolyn 'CC' Currie, Cody 'Jackfish' Jackson, Heather 'Yeet' Jovanovic, Krista 'Barney' Robertson, Lauren 'LT' Timlick, Leah 'Dickie' Dickenson, Lisa 'Tox Mom' Peters, Dr. Madeline 'Fun Stan' Stanley, and Nic 'Blandsy' Blandford. Your camaraderie, humour and craziness made the toughest moments manageable, even…. enjoyable. Love you pallys. To Dr. Vince Palace, from its inception through the field season, your leadership and adaptability were instrumental during the uncertainties of a global pandemic. Thank you for being both a golf partner and a mentor, ensuring we could "PROVE IT" and bring this project to fruition.

To my family and friends, your steady belief in my unconventional dream to "live in the bush collecting frogs" was extraordinary. Your faith and encouragement mean more than words can express.

Lastly, to the Experimental Lakes Area, as Fleetwood Mac so beautifully expressed:

"What are you doin', are you missin' me?
The way that I'm missin' you now?
The river goes on and on,
and the sea (lake) that divides us
is a temporary one."

Cheers to all the late nights, early mornings, sunburns, fake storms, and tipped canoes. You will forever hold a special place in my heart and I will see you again!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 - Overview

1.1 Thesis Overview

Acknowledgement of the lack of scientific literature on oil toxicity and fate in freshwater aquatic environments in Canada came in 2015 when the Royal Society of Canada review '*The Behaviour and Environmental Impacts of Crude Oil Released into Aqueous Environments'* identified the need for high-priority research in this area and identified key gaps (Lee *et al.*, 2015). In the context of freshwater, the 2015 Royal Society of Canada (RSC) report identified several Research Needs and Recommendations. Amongst them was the need for field studies (beginning with a simulated oil spill event) in varying freshwater ecosystems to understand the environmental impacts of industry spill response strategies, thus testing their efficacy (Lee *et al.*, 2015). The RSC report also identified the need for a priority-directed baseline monitoring on these poorly understood areas to then help comprehend research questions (Lee *et al.*, 2015, 2016). Further uncertainty of the best methods for oil spill clean-up in sensitive freshwater ecosystems and their environmental impacts was also a major concern and difficult to quantify and characterize due to the unique composition of the crude oil being transported in Canada. Uncertainties include the physical and chemical properties of oil, the behaviour of the oil in freshwater, and numerous biological sources (e.g. biotic and abiotic communities) all influence oil fate and toxicity in freshwater environments (Lee *et al.*, 2015). Another report, '*Spills of Diluted Bitumen From Pipelines: A Comparative Study of Environmental Fate, Effects, and Response'* by the National Academies of Sciences, Engineering, and Medicine; highlights in Section 4 other factors for consideration in pipeline oil spills, including spill conditions and location as well as the extent of the spill, the scope of the area and resources at risk (National Academies, 2016). To address some of these gaps, the International Institute for Sustainable Development Experimental Lakes Area and partners initiated in 2017 a research program composed of several interconnected studies. Amongst these, starting in 2018, the Freshwater Oil Spill Remediation Study (FOReSt) used *in-situ* enclosures to determine the effectiveness of non-invasive oil spill remediation techniques in different shoreline habitats of a boreal lake. Later, in 2021, the Floating Wetland Treatments to Enhance Remediation (FloWTER) study examined the efficacy of symbiotic plant-microbe relationships for *in-situ* degradation of oil-derived hydrocarbons via the use of engineered floating wetlands. This thesis examines the impacts on amphibians in the context of these two studies.

1.2 Crude Oil in Canada

Crude oil by definition, is a complex, naturally occurring liquid made of thousands of different hydrocarbon and non-hydrocarbon molecules (Lee *et al.*, 2015). Crude oil is characterized by its high density and high viscosity throughout the extraction, refining, and transport stages (Santos *et al.*, 2014; Marappa Gounder, 2019). Conventional crude oil, also known as liquid petroleum, is defined by Lee *et al.* (2015) as a mixture commonly transported in pipelines and is found in traditional oil wells. The transport of oil often can involve the use of diluents to reduce the viscosity of conventional heavy crude oils, allowing for more efficient transportation, particularly through pipelines (Kheraj, 2020). Crude oil provides substantial economic benefits to the Canadian economy; however, there are significant environmental risks, potential impacts, and concerns that are associated with the processes involved in the extraction, transport, and consumption of conventional heavy crude oil resources (Afzal *et al.*, 2019). This thesis focuses on the potential environmental issues related to transport (i.e., spills).

1.2.1 Physical and Chemical Properties

Crude oils have high acidity, high sulfur content, metal impurities, and mineral particles present, contributing to the complexity of the mixture (Lee *et al.*, 2015; Marappa Gounder, 2019). Crude oils have a unique composition of organic compounds that can be grouped into defined classes of chemicals, namely saturates, aromatics, resins and asphaltenes (SARA) (Lee *et al.*, 2015, 2016). Saturates typically make up a major component of liquid petroleum and are single non-polar hydrocarbons; aromatics are made up of one or more unsaturated ringed compounds (Santos *et al.*, 2014) and mono-aromatic hydrocarbons, commonly referred to as BTEX (benzene, toluene, ethylbenzene and xylenes) (Lee *et al.*, 2015). BTEX compounds are commonly present in crude oil (Lee *et al.*, 2015). Resins consist of high molecular weight compounds, whereas asphaltenes represent the most chemically complex, diverse, and highest molecular weight fraction of crude oil (Santos *et al.*, 2014; Lee *et al.*, 2015). The American Petroleum Institute (API) gravity scale is another characteristic that can be used to define and classify crude oils, and the API gravity is derived from the relative density (Santos *et al.*, 2014; Lee *et al.*, 2015). It is determined the API gravity of heavy crude oils is between 10°-22.3° API, while freshwater has an API gravity of 10°

API at 15 °C (Santos *et al.*, 2014; Lee *et al.*, 2015). Not only does the API gravity value provide indications on the oil's buoyancy, heavy crude oils typically between 10°-22.3° API, thus increasing the capacity to sink, additionally, API gravity also helps to determine the oil quality, refining and transport type (Santos *et al.*, 2014; Lee *et al.*, 2015).

1.2.1.1 Polycyclic Aromatic Compounds

Polycyclic aromatic compounds (PACs), including polycyclic aromatic hydrocarbons (PAHs), are naturally occurring contaminants released through natural processes such as wildfires and oil seeps (Fathallah *et al.*, 2012; Ball and Truskewycz, 2013; Bilodeau *et al.*, 2019). However, these can also be released through anthropogenic processes usually related to combustion processes, such as emissions from activities like mining, processing metals, vehicle exhaust, and petroleum refining and production (Ball and Truskewycz, 2013; Lee *et al.*, 2015; Bilodeau *et al.*, 2019). Conventional crude oils are comprised of smaller multi-ringed structures that are more resistant to biodegradation in the environment (Lee *et al.*, 2015). Although PAHs are made up of hundreds of compounds, 16 of the PAHs have been designated as high-priority pollutants by the United States Environmental Protection Agency (USEPA) (Hussar *et al.*, 2012; Lee *et al.*, 2015)(

Table 1-1). These include naphthalene, acenaphthylene, acenaphthene, fluorene, phenanthrene, anthracene (ANTH), fluoranthene (FLTH), pyrene (PYR), benzo[a]anthracene, chrysene, benzo[b]fluoranthene, benzo[k]fluoranthene, benzo[a]pyrene, benzo[g,h,i]perylene, indeno[1,2,3-c,d]pyrene, and dibenz[a,h]anthracene (Hussar *et al.*, 2012; Lee *et al.*, 2015; Balmer *et al.*, 2019). The reason 16 'priority pollutants' were chosen was because of their known toxic and carcinogenic effects to organisms within the environment relative to other PAHs (Hussar *et al.*, 2012; Lee *et al.*, 2015; Bilodeau *et al.*, 2019). These 16 'priority pollutants' are often used to characterize toxicity in environmental conditions and the ability to analyze these compounds (Hussar *et al.*, 2012; Balmer *et al.*, 2019). While PAHs are traditionally defined as hydrocarbons consisting of carbon and hydrogen arranged in multiple fused aromatic rings, the term PACs encompass a broader range of compounds and better reflects the complex mixtures found in the environment (Lee *et al.*, 2015).

Historical PAHs (i.e., the parent compounds) represent a small fraction of these crude oil mixtures (Lee *et al.*, 2015; Balmer *et al.*, 2019). PACs include not only the parent PAHs but also their alkylated derivatives, which are more prevalent in crude oil and environmental samples (Bilodeau *et al.*, 2019). While these 16 priority PAHs remain critical for regulatory purposes, recent research highlights the need to consider alkylated PAHs in freshwater environment toxicity studies (Bilodeau *et al.*, 2019).

1.2.2 Freshwater Oil Spills in North America

Canada transports and exports oil in several different ways, via rail, pipeline, coastal shipping, trucking, and this transport of oil is a concerning trend in Canada as it contains roughly 18% of the global freshwater within its borders, with 38% of those freshwater resources located near the southern, highly inhabited areas (Alsaadi *et al.*, 2018; Canada, 2020; Kheraj, 2020). The risks posed by oil spills are amplified by the vast pipeline network in Canada, which spans over 840,000 kilometres (Alsaadi *et al.*, 2018; Kheraj, 2020). In Alberta alone, there are over 426,000 kilometres of pipelines transporting roughly 2.9 million barrels of oil daily from oil sands operations (Alsaadi *et al.*, 2018; Canada, 2020; Kheraj, 2020). There are four different styles of a

pipeline, serving different purposes which makes up the extensive network throughout Canada, gathering pipelines, feeder pipelines, transmission pipelines and distribution pipelines collectively transporting 1.3 billion barrels per year (Alsaadi et al., 2018; Canada, 2020). Throughout the years, pipeline technologies, safety measures, and spill mitigation practices have developed with the objectives of improved economic efficiency and minimizing the percentage of oil product loss during spills (Kheraj, 2020). This, however, is negated by the fact that the occurrence, geographic density, and usage of oil pipelines have also increased, therefore causing an absolute volume of oil spills to also increase (Kheraj, 2020). For pipeline spills, the metric used by the Alberta Energy Regulator to characterize the rate of spills and performance is called Spill Rate (Canada, 2019). Spill rate is calculated as the volume of crude oil spilled (annual gallons) divided by the (total volume of mode transport (annual gallons) multiplied by the total distance the crude oil has moved (annual miles)(Canada, 2019). According to a 2019 report by The Railway Association of Canada, using data from Statistics Canada, the mean spill rate of pipelines over 15 years (2004-2018) was 18.1%, of 410,000 billion gallon-miles transported (Canada, 2019). Although spill rates have been declining since the early 2000s—from 47.7% spill rate of 19 billion gallon-miles transported in 2004 to 2.6% and 3.7% spill rates in 2013 and 2014, respectively—the total volume of spilled oil has continued to increase (Canada, 2019). For example, in 2018, as much as 281,000 litres of oil were spilt out of 44,000 billion gallon-miles transported via pipelines (Canada, 2019).

1.3 Oil Spill Remediation in Freshwater Environments

When assessing possible spill remediation efforts, it is important to examine a variety of factors when determining the best clean up methods, including oil type, location of the spill, time of year and weather, spill duration (chronic or acute exposure), and quantity of spilled oil (National Academies, 2016; Kheraj, 2020). In the event of an oil spill in freshwater environments, there are numerous conventional methods and standard industry practices to clean up and remediate sites affected by oil spills. However, typical industry practices use aggressive and invasive measures to clean up spills, which can harm sensitive shoreline habitats (Whelan et al., 2014; Palace et al., 2021). These measures may involve chemical dispersants, physical and chemical treatments, containment, and collection (Lee et al., 2015; Afzal et al., 2019).

Chemical dispersants are frequently used (currently not permitted in freshwater within Canada) to break down oil, making it easier for natural processes to degrade it; however, these are dependent on the spilt oils chemical and physical properties (National Academies, 2016). Physical and chemical treatments, such as washing shorelines (additionally not registered for use in freshwater in Canada) or using absorbent materials, aim to remove oil from affected areas, often deployed as an industry standard for remediation efforts (National Academies, 2016; Kharey et al., 2024). Containment efforts, including booms/curtains, are used to physically prevent the spread of oil and remove it from the water surface if deployed rapidly (National Academies, 2016). While effective to a certain extent, these methods often are ineffective in habitats such as wetlands and or flowing waters (National Academies, 2016). More recently, less invasive and more sustainable methods of remediation have also been identified, such as the use of bioremediation and natural processes (Lee *et al.*, 2015; National Academies, 2016; Afzal *et al.*, 2019). The methods implemented in this field experiment aim to utilize the natural processes of the environment to break down and remove contaminants rather than relying on more harmful and disruptive measures (Whelan *et al.*, 2014; Palace *et al.*, 2021).

1.3.1 Engineered Floating Wetlands

Engineered Floating Wetlands (EFWs), also commonly referred to as Floating Treatment Wetlands (FTWs), are a cost-effective and environmentally friendly approach to treating contaminated water, which involves using different types of aquatic plants (Rehman *et al.*, 2018). EFWs consist of a simple design that uses floating mats on which various macrophytes are cultivated with their roots suspended in the water column (Rehman *et al.*, 2019). Used in remediation efforts in sewage and wastewater, the planted vegetation within the EFWs provide an enriched/enlarged surface area for aquatic microbes to then grow on the roots as they are floating compared to natural rooted wetland plants, further enhancing the uptake of stressors and degradation of pollutants (Afzal *et al.*, 2019; Rehman *et al.*, 2019). These EFWs can use the natural properties of wetlands and floating vegetation, and EFWs are appealing for remediation efforts because they have low energy requirements and can be a sustainable alternative solution (Rehman *et al.*, 2018). Therefore, EFWs are a potentially viable solution when examining *in-situ* oil spill remediation in freshwater ecosystems, as they require little maintenance and minimal energy costs (Rehman *et al.*, 2018).

1.4 Toxicity and Environmental Risks of Oil Spills

1.4.1 Challenges of characterizing the toxicity of oils

A common approach to assessing the toxicity of the oil fractions that make their way into the water column is the use of a water-accommodated fraction (WAF) (Couillard *et al.*, 2005; Truter *et al.*, 2017). WAFs are defined as the 'hydrocarbons that will partition from oil to water during mixing processes including droplets' (Couillard *et al.*, 2005; Lee *et al.*, 2015). WAFs are formed in the field when crude oil is spilled and then mixes with freshwater, dispersing into the water column (Truter *et al.*, 2017). Thus, WAF is an important component to characterize effects for field experiments in freshwater environments (Couillard *et al.*, 2005; Truter *et al.*, 2017).

1.4.2 Susceptibility of Amphibians to Oil Spills

Amphibian populations have been steadily declining worldwide, and numerous environmental contaminants have been shown to play a role in this phenomenon (Blaustein *et al.*, 2003; Currie, 2018; Klemish *et al.*, 2018). Anthropogenic stressors, climate change, habitat degradation and alteration, increasing ultraviolet radiation, and pathogen introduction in aquatic and terrestrial environments are all possible drivers at regional and global scales (Alford and Richards, 1999; Blaustein *et al.*, 2003). Oil spills and oil-related contaminants could be a factor in local declines.

Amphibians can be affected by a wide array of stressors including, chemical pollution, such as pesticides or metals and due to amphibians sensitivity, low levels of the pollution in their environment can be harmful (Hopkins, 2007; Egea-Serrano *et al.*, 2012). An early standardized methodology was developed and is commonly used to assess alterations in endpoints such as mortality and growth called the Frog Embryo Teratogenesis Assay- Xenopus (FETAX) and was developed in 1983 (Sparling, 2010). Since then, studies characterizing the effects of process waters and multiple oil contaminants associated with oil and gas extraction have use FETAX (Sparling, 2010). More recently, the Larval Amphibian Growth and Development Assay (LAGDA) was validated to characterize the effects of environmental exposure of stressors on amphibians, using similar endpoints of growth and development through the early life stages as indicators (OECD,

2015). Among amphibians, *Xenopus* is also particularly prominent in laboratory research (Sparling, 2010; OECD, 2015).

Various species of *Lithobates* (e.g., *L. pipiens, L. catesbeiana, L. sylvaticus*) are commonly used in both laboratory experiments and field studies. Species such as wood frogs (*Lithobates sylvaticus*) are ideal for toxicological studies due to their high potential for bioaccumulation, making them especially useful for examining polycyclic aromatic compounds in water-accommodated fractions and oil spill scenarios (Moeun, 2018). Amphibians are often chosen for toxicology research because their permeable skin (dermal epithelia) causes them to be highly sensitive to environmental changes and contaminants, thus making them effective indicators of ecological health (Hopkins, 2007; Sparling, 2010; Ball and Truskewycz, 2013; Green *et al.*, 2013).

Early amphibian oil toxicity research found several factors influenced the toxicity of oils and impacts on frog health. The research characterized that the type of oil, the method of introducing the oil into the system, the weathering of oils and the exposure period collectively played a role in the LC_{50} values (Hedtke and Puglisi, 1982). They also reported that the LC_{50} values in frogs were less than those for freshwater fish (*Jordanella floridae*, *Pimephales promelas*) and wood frogs in flow-through tests were more sensitive than in static oil toxicity tests, perhaps indicating the reduced toxicity of oil when biodegradation and volatilization occurs (Hedtke and Puglisi, 1982). Lastly, the study found that emulsified oils were substantially more toxic than water-soluble fractions and surface oil (Hedtke and Puglisi, 1982).

A more recent field study conducted, Hersikorn and Smits (2011) looked at wetlands in the Athabasca Oil Sands and found that tadpoles raised in oil-sands process-affected materials (OSPM) areas complete metamorphosis significantly slower compared to tadpoles of control sites. Additionally, this study concluded that, interestingly, the body size of wood frog tadpoles was the largest in the tadpoles raised in young OSPM-affected wetlands when compared to reference sites and older OSPM-affected wetland sites (Hersikorn and Smits, 2011). This study contributed to the understanding of OSPM effects on wood frogs, but it recommends further research is needed to evaluate the long-term survivability of adult wood frogs from the study (Hersikorn and Smits, 2011). Although the compounds and toxicity of OSPW/ OSPM-affected areas are different from that of conventional heavy crude oil, extrapolation of the toxicity effect of similar compounds,

such as PAHs found within the OSPW/OSPM-affected areas remains relevant. Additionally, the methods in this study, in addition to the sampling endpoints examined, are similar to the methods and objectives identified for this thesis (Hersikorn and Smits, 2011).

The Royal Society of Canada Expert Panel recommended in section 4.1.5 that toxicity tests determine which characteristics of a species causes the sensitivity to oil, and recommends the classification of the most adverse exposure routes in these species (Lee *et al.*, 2015). For example, impacts of toxicants (PAHs) on amphibians have differing effects between species as well as at different developmental stages (Marquis *et al.*, 2006; Ball and Truskewycz, 2013). Ball and Truskewycz (2013) found that *Xenopus laevis* tadpoles seem to be resilient to PAH exposure "due to a poor hydrocarbon binding affinity to the AHR" (Ball and Truskewycz, 2013). Wood frogs experience delays in metamorphosis attributed to increased energy expenditure in the elimination of the toxicants (Hersikorn and Smits, 2011). A study published in 2000 examined oil-sands process-affected waters (OSPW) in an Alberta wetland with *Lithobates sylvatic*us present and also looked at endpoints, such as survival, growth, rate of development and deformities (Pollet and Bendell-Young, 2000). The study by Pollet and Bendell-Young (2000) found no wood frogs completed a full metamorphosis within wetlands around oil sands operations containing oil sands effluent, and although wood frog growth was not significantly affected, tadpoles experienced higher mortalities than *Bufo boreas* (Canadian toad) during exposures (Pollet and Bendell-Young, 2000). Within the last decade, a few publications have come out on the topic of the effects of PAHs and other PACs on amphibians Ball and Truskewycz (2013); Truter *et al.* (2017); Lara-Jacobo *et al.* (2019) but more specifically wood frogs (Currie, 2018; Bilodeau *et al.*, 2019; Denton, 2019; Mundy *et al.*, 2019; Patterson, 2019; Krohn *et al.*, 2020; Patterson *et al.*, 2022) to name a few.

A study using *Xenopus laevis* tadpoles highlighted the uncertainty of the effects of water-accommodated fractions of crude and diluted bitumen on thyroid systems of amphibians and aquatic vertebrates (Truter *et al.*, 2017). This study did, however determine that a crude oil WAF (25 g of crude oil/ Litre) disrupted thyroid signalling in response to 1% crude oil WAF (Truter *et al.*, 2017). In more recent years, only a handful of studies have examined oil effects now on amphibian behavioural and biochemical systems collectively. Many of these listed examples identified the need for further oil toxicity testing when looking at the exposure of crude oil and diluted bitumen on the effects on wood frog growth, survival, and development. There is little

research conducted on the effects of PAHs in field-collected amphibians and Sparling (2010) identified the need to further understand the differences in amphibian organic contaminant responsiveness (Sparling, 2010). The few studies available found that some amphibian embryos were insensitive to those chemicals containing PAHs at low levels, and it was also found that hatching success was not impaired in exposed green treefrogs (*Lithobates clamitans*) (Sparling, 2010).

1.4.3 Wood Frogs as Model Test Species

Presently, there are over 50 known species of amphibians in Canada, and amphibians are an abundant terrestrial vertebrate on the North American continent (Green *et al.*, 2013). Wood frogs (*Lithobates sylvaticus*) are the most widely distributed anuran species in North America and throughout much of their northern range are the only amphibian species present (Dodd, 2013; Green *et al.*, 2013). Both the Committee on the Status of Species at Risk in Ontario and the Committee on the Status of Endangered Wildlife in Canada have listed the status of the wood frog as Least Concern as of 2010 (Committee on the Status of Endangered Wildlife in Canada, 2010). Wood frogs undergo a metamorphic transition from aquatic larval stage to a terrestrial stage over a short duration and has been well documented in scientific literature (Taylor and Kollros, 1946; Gosner, 1960; Herreid Ii and Kinney, 1967; Hersikorn *et al.*, 2010; Navarro-Martín *et al.*, 2012; Ball and Truskewycz, 2013; Awkerman and Raimondo, 2018). Wood frogs, like many other amphibians, lay large egg masses in aquatic environments that hatch gilled tadpoles (Green *et al.*, 2013). Duration from the time eggs are laid until they hatch in the wild is also influenced by many factors such as water temperature, water quality, presence of predators (e.g., Dytiscidae family, dragonfly larvae from the *Anisoptera infra* order), conspecific competition and tadpole densities (Johansson and Nilsson, 1992; Dodd, 2013).

Wood frogs are useful biological indicators and are often implemented in toxicology studies because tadpoles and egg masses are unable to move away from environmental stressors present, thus making them more susceptible to direct exposure (Sparling, 2010). Then, once tadpoles complete metamorphosis, they are at risk of terrestrial contaminants throughout their habitats as well (Hopkins, 2007). An important aspect when considering freshwater oil spills is that wood

frogs inhabit many of the areas where pipelines are present in Canada and are commonly found throughout northwestern Ontario and are abundant at the study site (Experimental Lakes Area).

Using wood frogs as subjects, Mundy et al. (2019) quantified the concentration of PACs and PAC uptake in wood frogs present in the oil sands region of Alberta, Canada (Mundy et al., 2019). Although the study did not extrapolate and draw conclusions based on the results of PAC levels observed, it did discuss the potential for bioaccumulation of specific compounds in boreal food webs (Mundy et al., 2019). They also recommended the need for further research on the effects of the concentrations of PACs found within the tissues of the wood frog tadpoles (Mundy et al., 2019). Evidently, there is not enough evidence to extrapolate and generalize the effects of PAHs and contaminants across all amphibian classes Sparling (2010), which is why this amphibian field study and research will be scientifically appropriate.

1.5 Behavioural Endpoints in Environmental Toxicology

Behavioural assays are a useful sublethal tool in quantifying ecological effects because in amphibians, behaviour is often a primary indicator of exposure to an environmental stressor (Sparling *et al.*, 2010; Lee-Jenkins and Robinson, 2018). Peterson *et al.* (2017) also highlights this notion that due to the sensitive nature of behavioural changes, behaviour is a key indicators of exposure impact that can be observed at sublethal levels (Peterson *et al.*, 2017). However, Peterson *et al.* (2017) recognises the knowledge gap in the scientific literature in behavioural ecotoxicology. Understanding wood frog behavioural changes can help detect changes in tadpole foraging, sociality, activity and movement; therefore possibly altering tadpole predator avoidance capabilities (Bridges, 1997).

Video-tracking software is a tool in behavioural assays that is gaining popularity in amphibian tadpole research (*Rana temporaria* and *Lithobates sylvaticus*) (Denoël *et al.*, 2010; Sparling *et al.*, 2010). Video-tracking software was first introduced in the 1980s, and it has provided researchers with a more efficient way of quantifying behavioural data compared to more traditional manual counting methods (Henry *et al.*, 2019). Software such as EthoVision by Noldus or LiliTrack by Loligo has made computer-based analysis more commonplace, feasible and precise (Denoël *et al.*, 2010; Henry *et al.*, 2019). The use of behavioural analysis software is increasing in

research, with increasing publications of comparative studies examining behavioural methodology and analysis strategies (Denoël *et al.*, 2010; Henry *et al.*, 2019). More recently, behavioural endpoints are now credited, according to the US Environmental Protection Agency (USEPA), as an accepted method to measure amphibian fitness and exposure (Ågerstrand *et al.*, 2020). With the increasing usage of behavioural ecotoxicology, it is important to consider inter-individual variation in behaviour observations. Accounting for this variation is essential to quantify and accommodate for characterization effects (Careau and Garland, 2012; Bradley and Tracy, 2013). For example, in Denoël *et al.* (2010), the origin of egg mass was used as a random factor to address potential inter-individual variations (Denoël *et al.*, 2010). Additionally, behavioural parameters such as tracking activity may not reliably indicate survivability alone (Bridges, 1997). As Bridges (1997) noted, while increased speed during foraging raises the chance of predator detection, it also enhances likelihood of finding resources, balancing the overall impact on survivability (Bridges, 1997). Speed and distance are commonly used as indicators of an individual's ability to evade predators (Bridges, 1997). Lastly, the inability to detect and efficiently respond to the presence of predators or alarm signals from conspecifics increases the risk of predation (Polo-Cavia *et al.*, 2016; Bouffard, 2021).

1.5.1 Use of behavioural endpoints in amphibians

Not only are amphibians key indicators of environmental stressors, they are an abundant prey source for a range of aquatic and terrestrial predators therefore making them an important aspect in freshwater ecosystem energetics (Green *et al.*, 2013; Moeun, 2018). Behaviour quantification such as swimming ability, swimming patterns, activity and predator avoidance are all examples of indications used in past studies that link to higher-level responses. For example, a reduction in activity may lead to a delayed development and failure to reach metamorphosis, increasing vulnerability to predators (Bridges, 1997). Swimming patterns such as space use an important observation because it insinuates typical predator avoidance behaviours and individual "boldness" (Carlson and Langkilde, 2013). By looking at the percentage of time spent within the center of the behavioural arena in this study, border dwelling is an anxiety-like behaviour indicative of predator avoidance (Philibert *et al.*, 2016). Activity in the form of distance travelled is used to determine amphibian fitness, development and health. Finally, sociality (distance to nearest neighbour) is another indication of predatory avoidance behaviours of wood frog tadpoles

(Ågerstrand *et al.*, 2020). In a behavioural study conducted at the Experimental Lakes Area in 2019 on wood frog tadpoles, it was hypothesized the exposure to hydrocarbons from a diluted bitumen freshwater oil spill scenario would impair survival, growth, and development (Patterson *et al.*, 2022). It was determined that exposure to diluted bitumen on larval wood frog behaviour statistically impacted activity, shown through a reduction in tadpole movement however diluted bitumen did not impact other parameters such as sociality or space use (Patterson *et al.*, 2022). This current research project built upon the information and research findings while refining behavioural methodologies and endpoints to characterize effects of CHV exposure in a similar freshwater environment.

1.6 Research context for this thesis - Field Studies

The International Institute for Sustainable Development Experimental Lakes Area (IISD-ELA) in northwestern Ontario, Canada, sought to address the research needs outlined by the *Royal Society of Canada Expert Panel* with the initiation of two multi-faceted projects. These projects specifically address the Report's recommendations of 'assessment of the impact of spilled crude oil in high-risk and poorly understood areas such as Arctic, deep ocean, and freshwater ecosystems' and 'field studies (including the controlled release of oil) involving a spectrum of crude oil types in different ecosystems and conditions, and research at sites of previous oil spills in Canada' (Lee *et al.*, 2015).

The Freshwater Oil Spill Remediation Study (FOReSt) used in-lake enclosures to determine the effectiveness of non-invasive oil spill remediation techniques in different shoreline habitats. The overarching goal of the FOReSt was to a) 'determine the effectiveness of non-mechanical and biological methods for cleaning spilled oil from freshwater aquatic environments, and b) develop non-lethal and minimally invasive monitoring techniques to measure oil exposure and recovery in aquatic wildlife populations' (Palace *et al.*, 2021). Furthermore, the goal of the project was to examine the fate and behaviour of oil in biotic and abiotic communities within freshwater ecosystems (Palace *et al.*, 2021).

The second oil spill experiment, called the Floating Wetland Treatments to Enhance Remediation (FloWTER) study, examined the efficacy of symbiotic plant-microbe relationships

for *in-situ* degradation of oil-derived hydrocarbons. As part of a larger FLoWTER study, 26 mesocosms contained different ratios of aquatic plants within their engineered floating wetlands, in addition to treatments containing CHV-WAF. The objective of the study utilizing these 26 tanks was to develop specific protocols for optimizing the microbiome associated with EFWs to enhance the degradation of oil-derived hydrocarbons in freshwater environments.

1.6.1 Freshwater Oil Spill Remediation Study

The overall, multi-faceted FOReSt project conducted a model conventional heavy crude spill within shoreline enclosures and then monitored and measured varying environmental parameters throughout its duration. Physical, chemical, and biological oil spill remediation methods were compared for the main purpose of determining effective remediation measures for freshwater oil spill events.

In 2021, the goal of the amphibian portion of the multi-faceted FOReSt project was to characterize the effects of *in-situ* chronic exposure to CHV–derived hydrocarbons on the biochemistry, growth and development on wood frog (*Lithobates sylvaticus*) tadpoles. Developmental and morphometric measurements were combined with the assessment of bioaccumulation to determine the effects of oil-spill-impacted waters on the early life stage wood frogs. The experiment was conducted in six large-scale *in-situ* enclosures in shoreline environments (15 meters in length, 5 meters wide, with a maximum depth of 1.5–2 meters, and a capacity ranging from ~23,000 to 47,000 Litres) at a freshwater research site at the IISD Experimental Lakes Area (IISD-ELA) called Lake 260, (Kharey *et al.*, 2024). Three of the treatment enclosures contained 1.5kg of weathered Conventional Heavy crude oil (CHV) (n=3), obtained from pipeline stocks provided by the Canadian Association of Petroleum Producers (CAPP), while the remaining three enclosures served as un-oiled controls (n=3). Refer to Table 2-1 on quantities of oil that remained within the in-situ enclosures prior to primary cleanup for 72 hours, prior to shoreline washing and absorbent media retrieval of floating product only with no attempt to scrub shoreline materials or enclosures.

1.6.2 Floating Wetland Treatments to Enhance Remediation

The second study, Floating Wetland Treatments to Enhance Remediation (FloWTER) built upon the FOReSt project to assess the impacts of oil spills in freshwater ecosystems, while researching the capacity of engineered floating wetlands, naturally occurring wetland organisms and vegetation in the remediation of the conventional heavy crude (CHV) derived hydrocarbons of water-accommodated fraction (WAF) present as well as the duration of CHV- WAF degradation. Parameters measured include water chemistry, planted wetland vegetation growth within the engineered floating wetlands, as well as rhizome and biofilm development.

Within the mesocosms, differing ratios of planted vegetation were established on EFWs to enhance the capacity of wetland plants to metabolize substrates rather than to simulate natural wetland environments. Measurements on the effects of CHV-derived hydrocarbons in the WAF as well as the efficiency of wetland plants as remediation capacities, were examined. In addition to the validation of non-invasive biological methods, this research examined the potential impacts of oil spills on amphibians. Characterizing the impacts conventional heavy crude oil has on the traditional toxicity endpoints on wood frog tadpoles can assist in understanding how amphibian populations may respond in the event of an oil spill scenario. The amphibian component of the FloWTER study examined endpoints of exposure to CHV-derived hydrocarbons (WAF) on wood frog tadpole growth, development, and metamorphosis via traditional morphological methods as well as bioaccumulation approaches. Effects on behavioural responses and histology are also evaluated to further characterize effects in a multilevel quantification. Treatment mesocosms were organized as follows: three mesocosms contained water accommodated fraction and planted engineered floating wetlands (n=3); three mesocosms contained water accommodated fraction and unplanted engineered floating wetlands (n=3); and three mesocosms served as controls, containing no water accommodated fraction but with planted engineered floating wetlands (n=3). Engineered floating wetlands are constructed out of a buoyant material and are widely used in restoration efforts that use wetland plants for remediation and uptake of stressors (Rehman *et al.*, 2019). In my study, plant species, including *Typha* sp. (Cattail), *C. atherodes* (Slough Sedge), and *C. lasiocarpa* (Wiregrass Sedge) were included to determine the optimal ratio of plants that would contribute to the biological uptake of toxins and polycyclic aromatic compounds within the water accommodated fraction from the previously weathered conventional heavy crude oil (FOReSt -

Chapter 2). These EFWs within the mesocosms also provide a similar structure, habitat and refuge for early life stage amphibians found in natural environments. By incorporating EFWs into the experiment, tadpoles can sustain their natural grazing behaviour on aquatic macrophytes, detritus, leaf litter, and biofilm, eliminating the need for supplemental feeding.

1.7 Objectives and Hypotheses

Using *in-situ* enclosures in a boreal lake (a subproject of the FOReSt study- Chapter 2) , the **first objective** of my thesis was to characterize the effects of chronic exposure to conventional heavy crude oil–derived hydrocarbons on the growth, development, histopathology, and bioaccumulation in *Lithobates sylvaticus*. I hypothesize that:

H1: With the addition of CHV into the shoreline enclosures, growth, development, histopathology, and bioaccumulation parameters will be negatively affected in wood frog tadpoles.

H_0: There will be no significant differences in growth, development, histopathology, and bioaccumulation parameters in wood frogs resulting from chronic exposure to conventional heavy crude oil–derived hydrocarbons during larval stages.

Using mesocosms, the **specific objective** of my thesis, a subproject of the FLOWTER study- Chapter 3) is to characterize the effects of chronic exposure to CHV-derived hydrocarbons (WAF) on the growth, development, histopathology, and bioaccumulation of Lithobates sylvaticus. Additionally, to characterize the effects of chronic exposure to CHV-derived hydrocarbons (WAF) on the behaviour of *Lithobates sylvaticus*. Lastly, to characterize the influence of EFW presence/absence on ameliorating the growth, development, and behaviour responses of *Lithobates sylvaticus* chronically exposed to CHV-derived hydrocarbons (WAF). I hypothesize that:

H_1: With the addition of CHV-derived hydrocarbons (WAF) growth, development, histopathology, and bioaccumulation parameters will be negatively affected in wood frog tadpoles.

H_2: Tadpole activity will be negatively altered with exposure to CHV-derived hydrocarbons (WAF) therefore potentially affecting long-term survivability.

H_0: There will be no significant differences in growth, development, histopathology, bioaccumulation and behavioural parameters of CHV-derived hydrocarbons in wood frogs resulting from chronic exposure during larval stages.

Table 1-1 List of 16 Priority Polycyclic Aromatic Hydrocarbons designated by the US Environmental Protection Agency (Balmer *et al.*, 2019).

Priority Polycyclic Aromatic Hydrocarbons	CAS Number	Number of Rings	Molecular Weight
Naphthalene	91-20-3	2	128.17
Acenaphthylene	208-96-8	3	152.19
Acenaphthene	83-32-9	3	154.21
Fluorene	86-73-7	3	166.22
Phenanthrene	1985-01-08	3	178.23
Anthracene	120-12-7	3	178.23
Fluoranthene	206-44-0	4	202.26
Pyrene	129-00-0	4	202.26
Benzo[a]anthracene	56-55-3	4	228.29
Chrysene	218-01-9	4	228.29
Benzo[b]fluoranthene	205-99-2	5	253.32
Benzo[k]fluoranthene	207-08-9	5	253.32
Benzo[a]pyrene	50-32-8	5	253.32
Dibenzo[a,h]anthracene	215-58-7	6	278.35
Benzo[g,h,i]perylene	191-24-2	6	276.34
Indeno[1,2,3-cd]pyrene	193-39-5	6	276.34

Chapter 2 - Characterizing the Effects of Chronic Conventional Heavy Crude Oil Exposure on the Growth and Development of Larval Wood Frog (*Lithobates sylvaticus*) within Boreal Lake Littoral Enclosures of the FOReSt study.

ABSTRACT

Few studies have examined the effects of oil spills and their response methods on the early life stages of *Lithobates sylvaticus* under field experimental conditions. The Freshwater Oil Spill Remediation Study (FOReSt) was undertaken to determine the effectiveness of non-invasive oil spill remediation techniques on shoreline habitats and possible adverse ecological impacts at the field scale. Littoral enclosures (15 meters in length by 5 meters wide at a depth of 1.5-2 meters, ranging from ~23,000 to ~47,000 litres) were deployed on the shoreline of a lake at the International Institute for Sustainable Development-Experimental Lakes Area (Ontario, Canada) and assigned randomly to a reference treatment (n = 3) or a weathered conventional crude oil (CHV) spill (ranging from ~1372 grams to ~1653 grams, n = 3). The effects of the spill on the growth and development of wood frog (*Lithobates sylvaticus)* tadpoles were characterized over a 28-day exposure after a standard clean-up procedure of a shoreline rinse followed by collection of floating oil with absorbent pads. Bioaccumulation was observed with statistically significant increases in total polycyclic aromatic compounds (tPACs) in tadpole whole-body tissues in treated enclosures versus references. Despite this, no statistically significant treatment effects were found related to the growth or development of tadpoles. Qualitative histology of hepatic tissues found no apparent differences between treated and reference tadpoles. Overall, wood frog tadpoles did not exhibit ecologically significant mortality or negative effects on apical endpoints under this spill and response exposure scenario. We conclude that the standard spill clean-up procedures following the removal of oil after three days in this experimental design were adequate to reduce the impact of remaining low PAC levels and thus diminish the toxicological effects of PACs on wood frog tadpoles.

2.2 Introduction

Freshwater oil spills have become increasingly frequent over the past few decades, driven by the growing global reliance on oil resources for energy and transportation (Kheraj, 2020). Canada, a significant global exporter of oil, plays a central role in this trend, transporting oil across the country by rail, pipelines, coastal shipping, and trucking. The risks posed by oil spills are amplified by the vast oil pipeline network in Canada (Canada, 2020). Over the years, advancements in pipeline technologies, safety measures, and spill mitigation practices have aimed to enhance production and reduce losses from spills (Kheraj, 2020). However, the rising frequency, geographic concentration, and increased usage of oil pipelines have contributed to a higher overall volume of oil spills (Alsaadi *et al.*, 2018; Kheraj, 2020). When oil spills occur in freshwater environments, they could pose significant risks to those ecosystems. Polycyclic aromatic compounds (PACs), including polycyclic aromatic hydrocarbons (PAHs), are particularly concerning in the context of oil spills. These compounds are released during natural processes, such as wildfires and oil seeps, but are also produced by human activities, including petroleum refining, vehicle emissions, and oil extraction (Fathallah *et al.*, 2012; Ball and Truskewycz, 2013; Bilodeau *et al.*, 2019). In 2015, the Royal Society of Canada published a review that highlighted the lack of scientific understanding surrounding the fate and toxicity of oil in freshwater environments (Lee *et al.*, 2015). The report called for high-priority research to examine the relationships between oil exposure and its associated toxicity, with a particular focus on field studies in freshwater ecosystems. These studies are crucial for evaluating the effectiveness of industry spill response strategies and developing baseline monitoring protocols to improve environmental protection efforts (Lee *et al.*, 2016). Assessing effective spill remediation strategies requires considering several factors (National Academies, 2016; Kheraj, 2020). The complexity of crude oil, which is made up of thousands of different hydrocarbon and non-hydrocarbon molecules, presents unique challenges for remediation efforts (Lee *et al.*, 2015).

Understanding the interaction of crude oil with freshwater ecosystems remains a challenge, particularly given the variability in oil properties, environmental conditions, and biological factors. The effectiveness of spill response methods and the long-term impacts on freshwater habitats and organisms are difficult to quantify (National Academies, 2016). To address this, in 2021, the

Freshwater Oil Spill Remediation Study (FOReSt) aimed to fill knowledge gaps by evaluating the fate and effects of a simulated conventional heavy crude oil (CHV) spill that went through a standard clean-up procedure of shoreline rinse and collection of floating oil with absorbent pads. The FOReSt study was a multifaceted research project that measured numerous parameters and endpoints throughout the ecosystem, focusing on evaluating the effectiveness of non-invasive oil spill remediation techniques on shoreline habitats and their relative ecological impacts. The study was conducted at the International Institute for Sustainable Development Experimental Lakes Area (IISD-ELA) in northwestern Ontario, Canada.

Applying aspects of the larger FOReSt study, this thesis research focused on the amphibian component and aimed to assess the biochemical, growth, and developmental responses of wood frog tadpoles to *in-situ* chronic exposure to CHV. Following the recommendations of the Royal Society of Canada Expert Panel, which emphasized the need to identify species-specific sensitivities to oil and classify the most harmful exposure routes (Lee et al., 2015). For example, the impacts of toxicants (PAHs) on amphibians have differing effects between species as well as at different developmental stages (Marquis et al., 2006; Ball and Truskewycz, 2013). This project served to address some of the previously highlighted research gaps regarding oil spills in freshwater systems (Lee et al., 2015). To assess these effects, the study combined developmental and morphometric measurements with bioaccumulation analyses of total polycyclic aromatic compounds (tPACs) in tadpole tissues, alongside histological examinations of liver tissues, to determine the overall impact on early-stage wood frogs.

The main objective of this study was to characterize the effects of chronic exposure to conventional heavy crude oil on the growth and development, histopathological changes, and bioaccumulation in *Lithobates sylvaticus* within a freshwater ecosystem. The central hypothesis of the study was that the addition of CHV to shoreline enclosures would negatively affect the developmental parameters of wood frog tadpoles compared to shoreline enclosures without the addition of CHV. The null hypothesis posited no significant differences in developmental parameters resulting from chronic exposure to conventional heavy crude oil during the wood frog larval stages.

2.3 Materials and Methods

2.3.1 General Experimental Design

My study was conducted on Lake 260 (49°41'44.4" N 93°46'04.1" W), a 32.76-hectare freshwater research lake located in northwestern Ontario at the IISD-ELA (Figure 2-1). Lake 260 has a maximum depth of roughly 14 meters, is isolated from accessible waterways, and was selected for its consistency of shoreline and substrate (Figure 2-1). For this study, six shoreline enclosures were installed and constructed of impermeable plastic curtain (woven Novathene plastic; Curry Industries, Winnipeg, Manitoba, Canada) with a floating vinyl-wrapped Styrofoam 'boom' (Dow, Midland, Michigan, USA) at 5 meters wide by 15 meters long (Kharey *et al.*, 2024). The lengths of the enclosure were 10 meters, situated in the water while, the remaining 5 meters were located on the terrestrial shoreline (Kharey *et al.*, 2024). The enclosures were sealed to the underlying sediment with the curtain to prevent exchange into and out of the aquatic environment. This was done by laying down a double row of sandbags along the curtain-substrate interface and anchoring the floatation collar with aircraft cable at multiple points with rock anchors to withstand wave action. The enclosures reached a max depth of 1.5-2 meters and were in peat/organic substrate (Kharey *et al.*, 2024). Total volumes were greater than 20,000 L of water, and initially between 23,447 L and 46,603 L (SI Table 2-1). Please see Kharey *et al.* (2024) for further details on the study site.

The enclosures were installed in early June 2021 and then were left for a minimum of 7 days to allow sediment to settle before treatment. Treatments were assigned randomly to enclosures. On June 25th, 2021, three enclosures (n=3) were treated with pre-weathered (36-hour) conventional heavy crude (referred to as the oiled treatment enclosures), and three enclosures (n=3) were left untreated (referred to as not oiled reference enclosures) to function as reference systems (Kharey *et al.*, 2024).

Before addition to the enclosures, conventional heavy crude oil was weathered for 36 hours by pouring conventional heavy crude oil into stainless steel holding pans (1.1-meter diameter) containing lake water from Lake 260 (Kharey *et al.*, 2024). The weathered oil was collected using stainless steel instruments and transported in glass jars to the enclosures where ~1.5kg of

weathered CHV was applied directly to the water surface within 50cm of the shoreline. Then, four days after oil application, the initiation of conventional standard industry clean-up methods began to remove the CHV from the enclosures using low-pressure freshwater flushing and then oleophilic polypropylene sorbent pads and physical collection of the floating oil sheen (Kharey *et al.*, 2024). The mass of conventional heavy crude oil added into the treatment enclosure surface waters and shoreline was determined and calculated based on achieving a 0.1-cm oil thickness within the enclosure, or CHV approximately between 1,372 g and 1,653 g (mean of 1,477 g, n=3) (Palace *et al.*, 2021) (Table 2-1). Further methods regarding oil spills and cleanup of the enclosures are detailed in (Palace *et al.*, 2021; Kharey *et al.*, 2024).

2.3.2 Wood Frog Rearing and handling

All work with *Lithobates sylvaticus* (wood frog) was performed under an approved animal care protocol (University of Manitoba Animal Care Protocol- FOReSt F20-006-AC11576). Wood frog egg masses (n=7) were collected from Lake 227 at the IISD-ELA field station on May 4th, 2021, and transported back to a rearing area (SI Figure 2-1). This setup consisted of two 600-litre plastic cattle tubs that were filled ¾ with water from Lake 240, and then within each larger tank, up to seven separate 70-litre tubs, also filled ¾ with water from Lake 260, held the separate wood frog egg masses to prevent overcrowding. These animals were used for two simultaneously conducted studies (see also Chapter 3).

Temperature and basic water quality parameters (specific conductivity, pH, alkalinity, and dissolved oxygen levels) were monitored daily in the rearing tubs containing the wood frog egg masses. Water in the tubs was gently replaced with fresh water from Lake 260 at least every third day following recommended rearing procedures (Bridges, 1997). Water temperature during the rearing period ranged from 3.0°C to 21.7°C (mean of 12.5°C ± SD 5.2°C), while outside air temperature fluctuated from -1.6°C to 34.8°C (mean of 14.4°C ± SD 7.6°C) (SI Table 2-2). Hobo logger recorded temperatures located in (n=7) rearing tubs recorded a range of 2.7 °C - 30.2 °C (mean of 13.74 °C ± SD 4.9 °C).

Each egg mass was contained in a small hairnet to keep them afloat and self-contained. This allowed for the individual egg masses to be monitored daily for 11 days during the egg stage. The

mean hatching date of tadpoles from the egg mass was May 15, 2021 (± 3.5 days), determined by the emergence of embryos from the egg mass. Recently hatched tadpoles were fed frozen kale flakes during the pre-exposure holding time.

On June 15th, 2021, before moving the hatched tadpoles to the six *in-situ* enclosures, a subsample (n=20) of tadpoles were euthanized (using a buffered MS-222 solution), morphological endpoints were measured, and then tadpoles were frozen for pre-exposure baseline morphological determination. Tadpoles were reared in rearing tubs for a period from May 4th to June 15th, 2021. Pre-exposure measurements were recorded, indicating that the mean Gosner stage of the tadpoles was 25.3 ± SD 0.5. The mean snout-vent length was 10.9 mm ± SD 3.3 mm, while the mean total length was 27.9 mm ± SD 9.7 mm. Additionally, the tadpoles had a mean weight of 0.3 grams ± SD 0.3 grams. These baseline measurements were used to assess the initial development and size of the tadpoles prior to experimental treatment.

Any unhatched eggs were then removed from the rearing tubs. For timing feasibility, within the rearing period, all unhatched eggs and larvae mortalities experienced did not have parameters beyond the generalized Gosner stage measured. Any hatched tadpoles exhibiting abnormal behaviour or visible malformations were removed and euthanized. Malformations and normal and abnormal behaviour were defined by the FETAX Assay for Evaluation of Developmental Toxicity in combination with The Larval Amphibian Growth and Development Assay (Mouche et al., 2011; OECD, 2015).

Conventional heavy crude oil was added to the treatment enclosures on June 25th, 2021, and tadpoles were added seven days prior to allow tadpoles time to acclimate. Seven days prior to the addition of conventional heavy crude oil, tadpoles were transported to the in-lake enclosures in plastic bags (10 cm x 30 cm) containing water from their rearing tubs (originating from L260, the study lake) with a headspace of air (50% volume). Bags were left for roughly 30 minutes in the enclosures to acclimatize tadpoles to water temperature prior to release into their cages (19-litre cages with fine rubberized screens with top openings sealed by Velcro). A total of 14 early life stages (31 days old, mean Gosner stage 25.3 ± SD 0.4) wood frog tadpoles were placed in four cages in each of the six in-lake enclosures (SI Figure 2-2, Figure 2-1).

Each cage also contained a 2 - 7 cm pool noodle piece, which prevented the cages from entirely sinking within larger enclosures and allowed the tadpoles to rest as well as serve as an emergence area should any fully metamorphose during the study. Frogs in each cage were free-fed frozen kale every second day, and tadpoles were then monitored every second day throughout the duration of the study. Mortalities and tadpoles experiencing abnormalities and visible morphological malformation were removed and euthanized by overdose in buffered MS-222 solution (0.1 g/L MS-222 and 0.2 g/L sodium bicarbonate in lake water) (Robinson *et al.*, 2017). Malformations and normal and abnormal behaviour were defined by the FETAX Assay for Evaluation of Developmental Toxicity in combination with The Larval Amphibian Growth and Development Assay (Mouche et al., 2011; OECD, 2015).

Tadpoles were assessed daily and maintained up to the onset of metamorphosis (Gosner stage 40-42) at which point they were lethally sampled for morphometric data, and whole-body samples were prepared for either bioaccumulation or fixed for histology. Exposure time for the in-lake enclosures ranged from 7 to 21 study days (July 2nd to July 23rd, 2021) when tadpoles were removed and euthanized once they reached determined sampling timepoints denoted by Gosner stage. Target endpoints are identified in section *2.5.2 Apical Endpoints* below.

2.3.3 Water Chemistry

2.3.3.1 General Water Quality

Water within the *in-situ* enclosures was sampled throughout the experiment and analyzed by the IISD-ELA Chemistry Laboratory following standard procedures (Havens *et al.*, 2024). Water quality was assessed up to five times throughout the tadpole exposure period (SI Figure 2-3). Quantified parameters relevant to the tadpoles included alkalinity (Alk), ammonia (NH_3), calcium (Ca), chlorine (Cl_2), dissolved inorganic carbon (DIC), dissolved organic carbon (DOC), particulate carbon (PC), particulate phosphorus (PP), particulate nitrogen (PN), total dissolved phosphorus TDP), total dissolved nitrogen (TDN), total dissolved solids (TDS), total suspended solids (TSS) (OECD, 2015). Refer to Stanley (2024) for further enclosure condition descriptions.

Conditions in the enclosures were also measured using a handheld multiprobe (YSI Professional Series Plus, YSI QUATRO, ISD-ISD-DO-COND, 18C100206, Xylem, Yellow Springs, Ohio) for a total of five times throughout the exposure period (SI Figure 2-3). Each enclosure was also outfitted with a HOBO Pendant MX Temperature and Light Data Logger (*Onset Computer Corporation*) placed upright ~15 cm beneath the water surface and facing south prior to the addition of CHV and study day 0 (SI Figure 2-3). These temperature loggers within the *in-situ* enclosures collected data at 5-minute intervals for the duration of the exposure period. Refer to Kharey *et al.* (2024) for additional methods on the study site (Kharey *et al.*, 2024).

2.3.3.2 Water-Column Hydrocarbon Chemistry

Weathered conventional heavy crude oil was added to the treatment enclosures on June 25th, 2021 (study day 0). Sampling for 44 oil-derived hydrocarbon compounds (presented as total polycyclic aromatic compounds tPACs) was conducted at four different time points during the tadpole exposure period (SI Figure 2-3). This sampling schedule aimed to capture variations in PAC concentrations over the course of the exposure period (Stanley, 2024).

The analysis of PACs in enclosure water samples followed the methods developed by Idowu *et al.* (2018) and detailed in (Stanley, 2024). Sample preparation and extraction was completed at IISD-ELA during the 2021 field season, and the samples were then transported to the Center for Oil and Gas Research and Development Laboratory (COGRAD) at the University of Manitoba where PAC analysis was completed. The methods applied in the analysis allowed for the detection and quantification of numerous PACs, including polycyclic aromatic hydrocarbons (PAHs) and alkyl-PAHs (APAHs) (Idowu *et al.*, 2018). For methods on the study site and further enclosure oil chemistry methodology, please also refer to (Kharey *et al.*, 2024). For this study, PAC data will be summarised as the sum of all forty-four analysed PACs, i.e. total PACs (tPAC) (SI Table 2-3).

2.3.4 Tadpole Endpoints

2.3.4.1 Survivorship

Survivorship was determined at the end of the exposure period by calculating the total number of tadpoles initially added to the enclosures and subsequently adjusting this figure by subtracting the tadpoles collected for specific endpoints and pre-determined sampling events. Additionally, all documented mortalities were recorded throughout the study period. Given the nature of the field experiment, a category for unknown or unaccounted tadpoles was included to address instances where the fate of certain individuals was undetermined. The survivorship rate was then calculated using the formula:

$$Survivorship\ Rate\ (\%) = \left(\frac{Total\ Additions - \left(Known\ Mortalities + Unknown\ Fate\ (Total\ Samples - Known\ Mortalities)\right)}{Total\ Additions}\right) x\ 100$$

2.3.4.2 Apical Endpoints

The classification of the Gosner stage classes in this study followed the methods described and applied by Navarro-Martín *et al.* (2012) and was based on the original Gosner staging system developed by (Gosner, 1960). Apical endpoints (total length, snout-vent length, wet weight and Gosner stage) in tadpoles were measured at three sampling events in the study. These developmental categories provided useful timepoints for sampling that represented key stages in tadpole development, and the phases of the study which sampling would occur were determined to best align with the developmental stages.

The first sampling event, "pre-metamorphosis", for baseline determination of apical endpoints occurred at GS26, which is denoted by the formation of the operculum (Gosner, 1960; Hersikorn and Smits, 2011; Navarro-Martín *et al.*, 2012). At this point a subsample of 20 tadpoles were euthanized and sampled. The second stage, "pre-metamorphosis", took place during Gosner stage GS 34-36, when the majority of tadpoles have digit development on their hind limbs,

suggesting this is a midpoint stage in wood frog metamorphosis (Gosner, 1960; Reeves *et al.*, 2011; Navarro-Martín *et al.*, 2012). A total of 10 "pre-metamorphic" tadpoles in each enclosure were selected from the four cages, euthanized, and sampled at this time. The third measurement and the final exposure sampling time-point (called second-event) occurred just prior to "metamorphic climax", at Gosner stage 41 when front limb buds (initially the left front limb bud) are just emerging (Taylor and Kollros, 1946; Gosner, 1960; Navarro-Martín *et al.*, 2012). At this stage all remaining tadpoles that reached GS 40-42, tadpoles were collected, euthanized and sampled. In the instance where tadpoles did not reach this final Gosner stage prior to "metamorphic climax", the exposure period was then terminated at 28 days post oil addition to *in-situ* enclosures.

As tadpoles reached specified predetermined Gosner stage, they were removed from the enclosures and placed in one-litre clear plastic containers filled with associated enclosure water with 25-50% airspace, placed in a cooler and transported back to the laboratory. There, apical endpoints were recorded after euthanizing with a lethal dose of tricaine methanesulfonate (MS-222; 500 mg/L) in lake water, buffered to a neutral pH = 7 with sodium bicarbonate methanesulfonate (Grant, 2008; Iannella *et al.*, 2017).

Developmental stage was assessed via visible observation of each tadpole and using the staging table for anuran amphibians by Gosner (1960), to the approximate nearest Gosner stage. Once tadpoles were euthanized, individual tadpoles were measured for wet weight, total length to the nearest 1.0 millimeter (mm) (distance from nose to end of straightened tail, measured with tadpole on their ventral side) and snout to vent length (measured from the nose to vent of the tadpole) (Gosner, 1960; OECD, 2009). Wet weight was measured in grams using a Denver Instruments 0.001-gram digital scale (Denver Instrument P-403 400g x 0.001g, 115 V External Calibration).

Lengths (total and snout-to-vent) were measured using a digital caliper. Note that tadpoles with tail damage or missing sections, either due to natural causes or that occurred during sampling or transport, were excluded from the total length calculations. This exclusion was necessary because accurately measuring the total length (from snout to tip of tail) would not have been possible in such cases. However, the length from the snout to the vent (snout-vent length) was still recorded for these tadpoles.

2.3.4.3 Histology

A subset of tadpoles from the second sampling endpoint (GS 40-42) were collected for histological assessment. Originally a total of six tadpoles per enclosure were targeted for this endpoint; however, due to sampling constraints, allocated amounts were not always achieved. Final sampled numbers are as follows; for the reference enclosures, 5 tadpoles were sampled from enclosure 2 while enclosures 5 and 6 each had six tadpoles sampled. Within the oiled treatment group, 6 tadpoles were sampled from enclosure 1, 4 tadpoles from enclosure 3 used, and 2 tadpoles from enclosure 4.

2.3.4.3.1 Histological Sample Preparation

At the end of each exposure period, sampled tadpoles were removed and euthanized and those tadpoles partitioned for histological analysis were then given a mid-ventral incision through the body wall, placed in individually labelled 50 millilitre (mL) falcon tubes, and fixed whole in 10% neutral buffered formalin. Later, the formalin fixative was decanted from each sample and replaced with 70% ethanol. The tadpoles were then dissected under a stereoscope to remove the livers, which were placed in labelled histocassettes and stored in 70% ethanol until further processing. The ethanol was decanted and replaced twice to remove excess fixative from the tadpole tissues.

Liver samples were submitted to the University of Manitoba Histology Services Lab for paraffin embedding, and the blocks were sectioned and stained to confirm the methods before proceeding with the full batch of samples. The liver histology samples were dehydrated in graded alcohols, cleared in xylene, and infiltrated with molten paraffin using a Shandon Citadel 2000 Tissue Processor at the University of Manitoba Histology Services Laboratory. The samples were embedded in paraffin blocks in random orientation. The blocks were step-sectioned (5 millimetre thick) at three different levels approximately 100 microns apart, mounted on glass slides, and stained with Harris' hematoxylin and eosin.

2.3.4.3.2 Histological Assessment

Livers were assessed qualitatively for hepatic changes in the experimental treatments relative to the reference enclosure samples, including changes in hepatocyte size, nuclear abnormalities, staining properties (i.e., the degree of the eosinophilic (pink) or basophilic (blue) staining of the hepatocyte parenchyma), cytoplasmic vacuolation, or any other notable differences. For each individual tadpole, representative sections were examined across all three levels. Similar methods were employed in previous toxicity studies in fish (Alcaraz *et al.*, 2021a; Alcaraz *et al.*, 2021b; Alcaraz *et al.*, 2022; Colville *et al.*, 2022). Slides were analyzed using a Zeiss Axiostar Plus microscope. Photomicrographs were captured using a Lumenera INFINITY1-1M digital camera, and Teledyne Lumenera Infinity Analyze software (Version 7.0.2.930).

2.3.5 Bioaccumulation of Total Polycyclic Aromatic Compounds

Once measured and euthanized, wood frog tadpoles designated for total polycyclic aromatic compound (tPAC) bioaccumulation analysis were put individually into 50 mL Falcon tubes and immediately placed on dry ice to preserve their integrity and prevent any degradation of the tissues. Samples were transported to the Center for Oil and Gas Research and Development Laboratory (COGRAD) at the University of Manitoba and stored at -80°C. Tadpoles were then prepared following similar methods to those outlined in Bilodeau et al., (2019) and Mundy et al., (2019) and prior to bioaccumulation analysis, the entire gastrointestinal (GI) tract was removed from each tadpole ensuring the quantified PAC concentrations were that of the tadpole body and not from undigested food or sediment within the tadpole GI tract (Bilodeau et al., (2019).

2.3.5.1 Tissue Extraction

PAC quantification/extraction within tadpole body tissues followed the protocol of Xia *et al.* 2023 (Xia *et al.*, 2023). Individual tadpoles were extracted using a Precellys Evolution Micro-bead homogenizer. Each batch of 16 samples included a method blank, a freeze-dried *Mytilus edulis* tissue (SigmaAldrich, 2974a) reference material blank and reference material spiked with the recovery internal standard (RIS) for quality assurance and quality control purposes and 13 RIS spiked tadpole samples. The 20 µL of 1 ng/ µL RIS spike solution consisted of deuterated 16 US

Environmental Protection Agency (US EPA) priority PACs: d8-naphthalene, d8-acenaphthylene, d10-acenaphthene, d10-fluorene, d10-phenanthrene, d10-pyrene, d12-Benz(a)anthracene, d12-chrysene, d12-benzo(b)fluoranthene, d12-benzo(k)fluoranthene, d12-benzo(a)pyrene, d12-indeno(1,2,3-c,d)pyrene, d14-dibenzo(a,h)anthracene, and d14-benzo(g,h,i)perylene (Balmer *et al.*, 2019).

Spiked samples were then extracted with 8 mL dichloromethane (DCM) using the Precellys microbead homogenizer with 15 mL micro-bead extraction tubes containing 2.5 grams of the 2.8 mm zirconium oxide beads. The homogenate was then centrifuged (Bertin, Thermo Scientific Sorvall ST16 Refrigerated Centrifuge- Sorvall ST 16/16R) and the bottom DCM layer was transferred to a 60 mL test tube. The extraction was repeated two more times and the DCM extracts combined. The microbead homogenizer vial was rinsed twice with hexane, which was added to the final extract pool. Any water from the tissue was removed through addition of sodium sulfate directly to 60 mL test tube and shaken. The sodium sulfate was removed by filtering into a sterile 30 mL test tube through glass wool. The sample was reduced to 10 mL using an N-Evap nitrogen evaporator for lipid determination.

2.3.5.2 Lipid Determination and Gel Permeation Chromatography (GPC)

Lipid content was accounted for when measuring and analyzing the bioaccumulation of total polycyclic aromatic compounds. Lipid content is measured as percent lipid within the tadpole tissues and presented as nanogram per gram of tissue (ng/g). This was achieved by pipetting 500 µL of the homogenate sample extract (section 2.3.5.1) onto a pre-weighed aluminum boat; the solvent evaporates, leaving the lipid behind. The weigh boat is weighed again and then the final percent concentration was calculated.

To prepare sample extracts for Gel Permeation Chromatography (GPC), the volumes were reduced to 2.5 mL by nitrogen evaporation and samples were transferred into a 10 mL test tube with 2x DCM rinses. Final volumes were adjusted f to 5 mL. The sample extracts were applied to the GPC column (J2-scientific AccuPrep, Columbia, MO, USA). Column description and GPC run parameters are described in Xia et al. 2023 (Xia *et al.*, 2023). The fraction eluted from the column that contained the analytes of interest were collected in a 250 mL round bottom flask. The

clean extract was reduced to 200 µL using a Genevac Rocket evaporator (Ipswich, Suffolk, UK) and N-Evap. The final extract was transfer into a 2 ml amber gas chromatography vial with a 300 µL glass insert and spiked with 20 µL of 1 ng/ µL d-anthracene as the instrument performance internal standard (IPIS). Extracts were stored at 4°C prior to GC analysis.

2.3.5.3 Gas chromatography mass spectrometry/ mass spectrometry for Tadpole Tissues

This was set up specifically to process tissues as per methodology established for biota by (Xia et al., 2023). The instrument used for the native and d-labelled PAC quantification was an Agilent 7890 GC coupled to a 7000C triple quadrupole mass spectrophotometer fitted with an electron ionization (EI) source. Helium was the carrier gas with sample injection volumes of 1 µL injected by a PAL RSI 85 auto sampler. GC/MS/MS conditions and details of the multiple reaction monitoring precursor and product ion transitions for the 16 priority PACs that were analyzed in the tadpole tissue samples can be found in (Xia et al., 2023). Refer to SI Table 2-3 for list of PACs analysed in this methodology.

2.3.6 Statistical Analysis

All data analysis was conducted in R version 4.2.2 (R Core Team 2024). Project set up and basic data wrangling was conducted using the following packages: here v. 1.0.1(Müller, 2020), janitor v. 2.2.0 (Firke, 2023), knitr v. 1.43 (Xie, 2014, 2015, 2023), lme4 v. 1.1.34 (Bates *et al.*, 2023), multcomp v. 1.4.25 (Genz *et al.*, 2023), multcompView v. 0.1.10 (Graves *et al.*, 2024), patchwork v. 1.1.2 (Pedersen, 2022), rmarkdown v. 2.27 (Xie *et al.*, 2018; Xie *et al.*, 2020; Allaire *et al.*, 2023) and tidyverse v. 2.0.0 (Wickham *et al.*, 2019). Model outputs were formatted using broom.mixed v. 0.2.9.5 (Robinson *et al.*, 2023). All figures were created using ggplot2 from tidyverse v. 2.0.0 (Wickham *et al.*, 2019) with additional functions from patchwork v. 1.2.0 (Pedersen, 2022), scales v. 1.3.0 (Pedersen, 2022; Wickham and Seidel, 2022), and viridis v. 0.6.5 (Garnier, 2023).

2.3.6.1 General Water Quality

The analysis of general water quality parameters was initiated by testing for normality using the Shapiro-Wilk test, with p-values calculated for each treatment group. Following this, an F-test was performed to compare variances between the oiled and non-oiled groups, with the results categorized by their significance. Subsequently, a t-test was conducted to evaluate the mean differences between groups under the assumption of equal variances.

A multiple comparison test was applied to the general water quality data, comparing the two treatments. Specifically, Tukey's Honest Significant Difference (HSD) test was used to compare the oiled and not oiled enclosure treatments and obtain p-values for visualizing differences. Since the mixed-effects model does not generate p-values, Tukey's HSD allowed for statistical comparisons between treatment groups.

2.3.6.2 Conventional Heavy Crude Oil Water Chemistry

To assess the average exposure of the tadpoles to oil-derived hydrocarbons throughout the exposure period, water-column tPAC data from the various analytical sampling events was used to calculate Time Weighted Average (TWA) concentrations for each enclosure. Due to the complexity of the study, it was not possible to pair sampling for tPAC with tadpole sampling. As such, the two sampling events for tadpole endpoints fell in between tPAC sampling days. To estimate the TWA concentration at the time of the two tadpole sampling events, a LOESS model was fit to the TWA concentration data for the tPAC analytical sampling days using the loess function from the basic stats package leaving all default settings. TWA concentrations for the tadpole sampling events in each enclosure were then predicted from these models with the predict function from the stats package.

Statistical methods applied for analyzing enclosure CHV water chemistry started with a Shapiro-Wilk test to assess normality, where p-values were calculated for each treatment group. Following this, an F-test was conducted to compare variances between the oiled and non-oiled groups, with the results categorized by their statistical significance. A t-test was then performed to evaluate mean differences between the groups, assuming equal variances. To evaluate the effects of treatments on the analysis of enclosure CHV water chemistry, a one-way ANOVA was then

performed for the enclosure CHV water chemistry. Next, a multiple comparison test was performed on the enclosure CHV water chemistry data. Specifically, Tukey's Honest Significant Difference (HSD) test was employed to compare the oiled and not oiled enclosures, enabling visualization of p-values since the mixed-effects model does not generate them.

2.3.6.3 Tadpole Endpoints

Tadpole apical endpoints were assessed individually and for each of the sampling events separately. To be able to use individual tadpole data (i.e. avoid condensing to one value per enclosure), without incurring on pseudo-replication, tadpole biological endpoints were analysed via a mixed effects model approach with the presence or not of oil as the main treatment variable and enclosure as a random effect. The lmer function within the lme4 package (Bates *et al.*, 2015) was employed to fit these models using the restricted maximum likelihood (REML) option. Conformity to model assumptions was assessed by visual inspection of the distribution of residuals. Pairwise multiple comparisons of the treatments was conducted using the lsmeans function from the emmeans package Lenth (2024) followed by the cld function from the multcomp package (Hothorn *et al.*, 2008).

2.3.6.4 Bioaccumulation of Total Polycyclic Aromatic Compounds

Tissue tPAC data was estimated in pools of tadpole tissue due to analytical method requirements. Several pools were available, however, for each enclosure. To be able to use the data from each pool directly (i.e. avoid condensing to one value per enclosure), without incurring on pseudo-replication, the tissue tPAC data was analysed in the same manner as described above for the tadpole endpoints using a mixed effects model approach with enclosure as a random effect. Pairwise multiple comparisons via lsmeans Bates *et al.* (2015) followed as described above to determine specific differences between treatments.

2.4 Results

2.4.1 Water Chemistry

2.4.1.1 Wood Frog Rearing

Pre-exposure measurements from the rearing tubs provided baseline conditions, and confirmed that the rearing conditions were suitable for wood frog tadpoles aligning with typical OECD guidelines (OECD, 2009, 2015)(SI Table 2-2).

2.4.1.2 General Water Quality

High-frequency temperature data recorded in the mesocosms from June 18 to July 23, 2021, showed no significant temperature (°C) differences among treatments as indicated by a two-sample Student's t-test ($t = 0.14$, $p = 0.89$). The oiled enclosures had a mean temperature of 23.2°C (± SD 2.9°C) and the not oiled enclosures had a mean temperature of 23.3°C (± SD 3.1°C) (SI Table 2-4, SI Table 2-5, SI Table 2-6, SI Table 2-7). Measurements using a YSI probe for dissolved oxygen (DO), pH, specific conductivity, and temperature across treatments from June 18 to July 23, 2021, a two-sample Student's t-test indicated no significant differences across treatments (SI Table 2-8, SI Table 2-9, SI Table 2-10, SI Table 2-11).

Water nutrient parameters were measured throughout June 18 to July 23, 2021, and a two sample students t-test indicated no statistically significant differences between the treatments (SI Table 2-15). The conditions in the enclosures were deemed acceptable for housing wood frog tadpoles until the sampling endpoints occurred (OECD, 2009, 2015).

2.4.1.3 Water-column Hydrocarbon Chemistry

The temporal distribution of water-column tPAC concentrations is presented in Figure 2-3. A clear increase on tPAC can be seen starting on day 0 (oil application day) and peaking at the end of the first week. Time weighted average (TWA) concentrations of total polycyclic aromatic compounds (tPACs) were calculated for the average exposure day associated to each of the main

sampling events. Significant differences in TWA concentrations of tPACs across the treatments, not oiled (reference) and oiled treatments across both exposure days (Table 2-2, Figure 2-3, Figure 2-4). Elevated levels of retene were detected in the TWA concentrations within both the oiled and unoiled (reference) enclosures when looking at the individual compounds, according to the findings in Stanley (2024). This complicates the interpretation of tPAC exposure, as it suggests that some compounds may originate from past environmental conditions rather than entirely from the present study.

2.4.2 Tadpole Endpoints

2.4.2.1 Survivorship

The mean survivorship for the oiled enclosures (83.1% ± SD 6.1%) was lower compared to the not oiled enclosures (87.8% ± SD 4.7%). The oiled group also had a higher mean unknown fate (6.3 tadpoles ± SD 0.3) compared to the not oiled enclosures (4.0 tadpoles ± SD 0.0). This shows us that the conditions did not cause high mortality events within the enclosures with only marginally lower survivorship in the oiled group. While the difference in known mortality between treatments was small, the oiled enclosures had slightly higher mean known mortalities (3.7 tadpoles ± SD 4.6) than the not oiled enclosures (3.0 tadpoles ± SD NA)(Table 2-3).

2.4.2.2 Apical Endpoints

The mean and standard deviation of all apical endpoints (weight, Gosner stage, snout-vent length, and total length) are shown in (Table 2-4). There were no significant differences ($p > 0.05$) for these endpoints between the oiled and not oiled treatments across both sampling events (SI Table 2-18, SI Table 2-19, Figure 2-5). Overall, the F-statistics indicate that there are differences in variability among the different endpoints measured in each GS class. However, none of the F-statistics suggest strong evidence of significant differences, as they are all low (SI Table 2-19).

2.4.2.3 Histology

There was no apparent difference between reference (n=17) and oiled treatments (n= 12) in the histological appearance of the liver (Figure 2-6). The liver parenchyma was comprised of relatively large and polygonal hepatocytes, with abundant cytoplasmic vacuolation resulting in eccentric nuclei. Cords of hepatocytes were interspersed with sinusoids and blood vessels containing erythrocytes (Figure 2-6).

2.4.2.4 Bioaccumulation of Total Polycyclic Aromatic Compounds

Examination of total polycyclic aromatic compounds within whole-body tadpole tissues found tissue tPAC concentrations to be statistically higher in tadpoles from the oiled treatment (mixed effects model approach followed by multiple comparisons). The model was fitted with as random intercepts model with enclosure as the random effect (sintaxt: lipid-corrected tissue concentration ~ treatment + (1|enclosure)). For the first sampling event, the mean concentration of tPACs in not oiled tadpoles was 1.03 ± SD 0.65 ng/g lipid. In oiled tadpoles, the mean concentration was significantly higher at 5.72 ± SD 1.90 ng/g lipid (SI Table 2-21). In the first sampling event (GS 34 – 36), significant differences were observed between oiled and not oiled treatments for both alkylated PAHs and tPACs, with oiled treatments having higher concentrations. No significant difference was found for parent PAHs.

For the second sampling event (GS 40-42), the mean concentration of tPACs in tadpoles within the reference (not oiled) enclosures was 0.26 ± SD 0.23 ng/g lipid, whereas in tadpoles within treatment (oiled) enclosures, the mean concentration increased to 2.36 ± SD 1.50 ng/g lipid (Table 2-5). Despite the ~10x difference in tPAC, these were not significantly different (Estimate = -2.096, P-Value = 0.071). No statistically significant differences were found in alkylated PAHs, tPACs, and parent PAHs (SI Table 2-20, SI Table 2-21, Figure 2-7) when analyzed individually.

2.5 Discussion

The objective of this study was to characterize the growth and development of wood frog (*Lithobates sylvaticus*) tadpoles over a 28-day exposure period following a simulated oil spill of

conventional heavy crude oil (CHV) in a boreal lake, which included a remediation step involving non-invasive treatment approaches. It was hypothesized that the addition of CHV to shoreline enclosures would negatively affect key developmental parameters in wood frog tadpoles. In addition to developmental measurements, the bioaccumulation of tPACs (total polycyclic aromatic compounds) and liver histology were analyzed to further characterize exposure and potential adverse effects.

Within the enclosures, water quality parameters revealed no significant differences between the oiled (treatment) and not oiled (reference) enclosures. The temperature remained stable and comparable across treatments, and nutrient parameters consistently indicated no significant impact of the conventional heavy crude oil on water nutrient parameters during the exposure period. Lastly, the measured water quality variables suggest that oil exposure did not appear to affect these basic water variables, and differences in water quality parameters are unlikely to be the drivers of any observed differences in the tadpoles.

As expected, the concentration of tPACs in the water column increased sharply in the treated enclosures following application and decreased over time, with lower values in the second sampling event compared to the first. Analysis of the fate of individual PACs during the study is presented; however, it is important to note that elevated levels of Retene, a priority pollutant identified by the United States Environmental Protection Agency (USEPA), were detected in the TWA concentrations in both the treatment and reference enclosures (Hussar *et al.*, 2012; Lee *et al.*, 2015). It is likely that the presence of past wildfires in the area contributed to these elevated baseline concentrations of Retene, complicating the interpretation of tPAC exposure (Stanley, 2024).

In characterizing the effects of conventional heavy crude oil on tadpole development, several apical endpoints were assessed, including Gosner stage, snout-vent length, total length, and weight, across two sampling events. The analysis included two sampling events at Gosner stages approximately 34-36 and Gosner stage 40-42 and determined that no statistically significant differences between treatments were found for these endpoints. These findings are not consistent with the scientific literature on polycyclic aromatic compound toxicity in amphibians, for the studies current PAC concentration levels up to maximum of 1,130.42 ng/L time weighted average

of tPACs. For example, one study found that *Xenopus laevis* exposed to coal tar sealants or PAHs, such as benzo(a)pyrene, experienced lowered hatching success, reduced growth and delayed development and metamorphosis (Bryer *et al.*, 2006). Similarly, field studies conducted by Hersikorn *et al.* (2010); (Hersikorn and Smits, 2011); Patterson *et al.* (2022) documented delays in wood frog development, increased deformities and altered thyroid hormone levels after oil exposure, pointing to compromised metamorphosis. Another field study by Pollet and Bendell-Young (2000) found no wood frogs completed a full metamorphosis when exposed to oil sands process-affected water (PAHs), growth was affected by oil sands process-affected water, and wood frog tadpoles experienced higher mortalities than *Bufo boreas* (Canadian toad) during exposures (Pollet and Bendell-Young, 2000).

The study by Hersikorn and Smits (2011) provided some reasoning as to the possible causes of those increased sizes, which included lower population density, genetic advantages allowing survival under toxicity, and lastly, age at the metamorphic climax, allowing more time for growth. When also examining the apical endpoints measured in a similar study conducted at IISD-ELA in 2019 (Patterson *et al.*, 2022), that study found wood frog tadpole total lengths ranged from 43.2-51.3 mm and 0.88–1.25 g for tadpole weight. The total length and weight of the tadpoles across the two treatments during both sampling events in this current study were greater compared to the Patterson *et al.* (2022) study. In the first sampling event, using the lower end of Patterson *et al.* (2022) the measurement range, tadpoles in the non-oiled enclosures of this study were approximately 9.8 mm longer and 0.62 g heavier, while those in the oiled enclosures were 12.4 mm longer and 0.72 g heavier. For the second sampling event, using the upper end of Patterson *et al.* (2022) range, tadpoles in the non-oiled enclosures were 9.2 mm longer and 0.65 g heavier, while those in the oiled enclosures were 8.6 mm longer and 0.75 g heavier.

To investigate the factors contributing to the increase in size observed in the current experiment, the enclosure density between the treatment and reference groups, as well as the mean age of the tadpoles at each sampling time point, were examined to further characterize potential effects (Figure 2-3, Figure 2-4). Firstly, the survivorship of tadpoles within enclosure cages only varied slightly between the non-oiled and oiled enclosures. The survival rates within both oiled and non-oiled enclosures were similar, with only marginally lower survivorship in the oiled group. Lara-Jacobo et al. (2019) used chemically enhanced WAFs with much higher tPAC levels

(152,000-247,000 ng/L TPACs), resulting in high mortalities (100%). When tadpoles were exposed to lower concentrations (700, 3,800–8,200 ng/L), the concentrations resulted in survival rates of 96%, 92% and 91.5%, respectively, demonstrating acute toxicity at these levels (Lara-Jacobo et al., 2019). This enclosure study, using similar tPAC levels (maximum at day 7 of 1,130.42 ng/L), recorded survival rates from 83.1% ± SD 6.1% to 87.8% ± SD 4.7%, suggesting wood frog tadpoles exhibit very minimal reduction in tolerance to CHV at the exposure concentrations, compared to the higher concentrations of WAF exposed to in the Lara-Jacobo et al. (2019). Thus, the presence of CHV appears to have no effect on wood frog tadpole survival throughout the study.

The numerous non-statistically significant findings do align with the summary from Sparling (2010) that revealed a few studies found some amphibian embryos were generally insensitive to PAHs at low levels (Sparling, 2010)(Patterson *et al.*, 2022). The observed lack of histological differences between treatments suggests that exposure to CHV did not cause detectable cellular damage or alterations in liver structure in the wood frog tadpoles within the *in-situ* enclosures. Within the enclosures, no apparent difference was noted between the reference group (n=17) and the oiled group (n=12) in the histological appearance of the liver (Figure 2-6). In both groups, the liver parenchyma consisted of large, polygonal hepatocytes with abundant cytoplasmic vacuolation, leading to eccentric nuclei. Research by Regnault *et al.* (2016); (2018) showed that prolonged BaP as low was 50 ng/L^{-1} (one of the 44 Polycyclic aromatic compounds (PAC) measured for analysis, SI Table 2-3) exposure in *Xenopus* tadpoles resulted in metabolic disruptions, such as liver damage (Regnault *et al.*, 2016; 2018). Additionally, Regnault *et al.* (2016); (2018) results showed an increase in lipid droplet content within hepatocytes, along with structural alterations in the parenchyma and irregular cell shapes (Regnault *et al.*, 2016; 2018). Hepatocyte cords were interspersed with sinusoids and blood vessels, which contained erythrocytes, indicative of normal histological architecture across both treatments. This can be compared to increased lipid droplets in liver cells and decreased pigment content when conducting histological structures of the liver, indicative of hepatic stress and have been documented in western clawed frog tadpoles exposed to BaP (Regnault *et al.*, 2014). The lack of liver abnormalities in this study's CHV-exposed tadpoles shows that CHV at the exposure level did not produce comparable endocrine-disrupting or metabolic effects at detectable levels in wood frogs (Regnault *et al.*, 2016; 2018). This consistent liver morphology across treatments implies that the

CHV exposure may not have led to significant liver damage, at least under the treatment conditions and exposure duration. If changes were to be observed, this likely could have been attributed to stress, thus, causing damage to tissues, an impact from PAC toxicity (Thambirajah *et al.*, 2019). This is affirmed by previous work done that wood frog tadpoles have efficient physiological mechanisms for polycyclic aromatic compounds, potentially conferring a degree of resilience to environmental oil contamination (Bilodeau et al., 2019). Moreover, liver histological assessments revealed no cellular damage or significant morphological changes between reference and treatment groups, further supporting the conclusion that CHV exposure did not cause detectable harm at the cellular level and contrast with existing literature on the toxicological impacts of oil-related contaminants, such as PAHs, on amphibian liver histology and function.

In research by Mundy *et al.* (2019), wood frogs were used as subjects with the aim of quantifying the concentrations of PACs and PAC uptake in frogs from the oil sands region of Alberta, Canada. Although the study did not derive conclusions from the observed PAC levels, it highlighted the potential for bioaccumulation within boreal food webs (Mundy *et al.*, 2019). Still, comparing this research to the current work is not precise as that study did not lipid-normalize wood frog samples or account for lipid correction, an approach included in this study (Mundy *et al.*, 2019). Analysis of tPAC concentrations in the tadpole tissue showed clear signs of bioaccumulation after 21 days in this enclosure study, with a statistically significant increase in tPAC concentrations in tadpole whole-body tissues in oiled (treatment) enclosures compared to the references in the first sampling event only. This indicates a greater exposure to tPACs leads to an increased uptake of these compounds at statistically significant concentrations. The study by Hersikorn and Smits (2011) found that the body size of wood frogs tadpoles was largest in the tadpoles raised in newer oil sand process-affected material (OSPM) wetlands when compared to reference sites (without OSPM) and older OSPM-affected wetland sites (Hersikorn and Smits, 2011). Although not comparable as the apical endpoints were not statistically significant, the bioaccumulation within the treatment tadpoles of this experiment showed that tadpoles demonstrated more significant bioaccumulation of CHV when concentrations (shown by time-weighted average) were greater. Several factors may explain this result. Hedtke and Puglisi (1982) reported that wood frogs in flow-through tests were more sensitive than in static oil toxicity tests, perhaps indicating the reduced toxicity of oil when biodegradation and volatilization occurs (Hedtke and Puglisi, 1982). Secondly, during the second sampling timepoint, tadpoles experienced

lower concentrations of tPACs within the water in the treatment enclosures, and thus, exposure concentrations were not consistent with the first sampling event (Figure 2-4). Lastly, these findings align with research by Bilodeau *et al.* (2019) which examined the elimination and bioaccumulation of PACs in anuran larvae, after a 24-hour exposure. Bilodeau *et al.* (2019) concluded that wood frog tadpoles are efficient at metabolizing and eliminating PACs as rapidly as 12 hours of exposure (Bilodeau et al., 2019). tPAC tissue concentrations were significant during the first sampling event and concentrations decreased in the second sampling event, likely due to both metabolic processes and reduced CHV exposure. Given that the exposure period in this experiment spanned 21 days, this could explain why significant bioaccumulation was observed during the first sampling event, when tadpoles were smaller at earlier Gosner stages and had shorter lengths, but not in the second sampling event. The accumulation of tPACs in tadpole tissues may not impact the developmental markers and apical endpoints of wood frog tadpoles but may have broader implications for predators and the overall food web in affected areas. Overall, wood frog tadpoles did not exhibit ecologically significant mortality or negative effects on apical endpoints under this spill and response exposure scenario. We conclude that the standard spill clean-up procedures following the removal of oil after three days in this experimental design were adequate to reduce the impact of remaining PACs and thus diminish the toxicological effects of PACs on wood frog tadpoles.

2.6 Conclusions

This study aimed to assess the effects of chronic exposure to conventional heavy crude oil (CHV) on the growth and development of wood frog (*Lithobates sylvaticus*) tadpoles, from cellular-level impacts to whole-organism developmental endpoints. When characterizing the effects of CHV exposure on wood frog tadpoles, it was determined that at the given concentrations within the treatment enclosures, wood frog tadpoles were not adversely affected by the tPACs in the CHV. Based on the experimental data, it was determined that the concentrations of CHV used in the treatment enclosures generally did not induce statistically significant adverse effects on wood frog tadpoles over the 21-day exposure period. Although bioaccumulation of tPACs was evident in the first sampling event, with a statistically significant increase in tPAC concentrations

in tadpole whole-body tissues in oiled (treatment) enclosures compared to the references, this trend was not consistently observed across all parameters measured.

This finding suggests that, as a whole organism in general, CHV exposure did not broadly affect all aspects of tadpole growth and development within the exposure period. Thus, the hypothesis that "the addition of CHV to shoreline enclosures would negatively affect developmental parameters in wood frog tadpoles" is rejected. The null hypothesis, which posits that there would be no significant differences in the developmental parameters of tadpoles exposed to CHV compared to references, is accepted. These results trend towards the notion that wood frog populations in freshwater environments exposed to similar CHV concentrations may maintain normal development and survivorship during their early life stages. Extrapolation and modelling of this non-significant effect and ecological impact from these results could potentially help characterize effects throughout other trophic levels, understanding toxic effects and concentrations of the listed priority PACs for wood frog tadpole predators. While bioaccumulation of tPACs was observed in the tadpoles, this was primarily significant during the first sampling event. The tadpoles appeared capable of metabolizing and potentially eliminating tPACs over time. The resilience of early life-stage wood frogs to effectively metabolize or not be negatively affected by tPACs in various capacities, from histological levels to bioaccumulation and apical parameters, is of ecological significance for wood frog and amphibian conservation.

Importantly, the findings of this study contribute to a broader understanding of how amphibians, particularly wood frog tadpoles, respond to freshwater oil spill scenarios. Although significant bioaccumulation of tPACs occurred early in the exposure period, the lack of adverse effects on growth, development, and liver histology indicates that early-life stage wood frogs may possess the resilience to CHV contamination at the tested concentrations. In an ecological context, this result may suggest that wood frogs and potentially other amphibian species may not experience severe developmental impairments in environments with moderate levels of CHV contamination. These results also have broader implications for ecological risk assessments of oil spills in boreal freshwater ecosystems. The ability of wood frog tadpoles to metabolize and tolerate tPACs could have cascading effects on the broader food web, particularly for the host of predators that rely on amphibians as a food source. Future research should further investigate the long-term effects of CHV exposure across different trophic levels and across different wood frog life stages.

In conclusion, this study highlights the importance of using amphibian models, such as wood frog tadpoles, to evaluate the ecological impacts of freshwater oil contamination. The findings provide valuable insight into the potential resilience of amphibian populations in the face of freshwater oil spills while also addressing the gaps in the current literature.

Table 2-1 Mass of weathered (36-hour) conventional heavy crude oil (CHV) applied to treatment enclosures at Lake 260 at the IISD-Experimental Lakes Area on June 25, 2021. After a 4-day weathering period, enclosures treated with CHV received industry-standard physical remediation methods, including shoreline cleaning. *Primary recovery was delayed due to wildfire in the area preventing primary recovery from occurring on a single day.

Enclosure	Date Applied	Weight Applied (grams)	Carbon Weight (grams)	Primary Recovery Date	Total Recovered (grams)	Per. Recovered (%)	Weight Remaining (grams)
1	25-June-2021	1,407	1,224.09	29-June-2021	202.50	14.39	1,204.50
3	25-June-2021	1,653	1,438.11	30-June-2021	356.40	21.56	1,296.60
4	25-June-2021	1,372	1,193.64	29-June-2021 & 30-June-2021*	508.60	37.07	863.40

Table 2-2 Time weighted average (TWA) concentrations of total polycyclic aromatic compounds (tPACs) from CHV oil sampled from enclosure water at the IISD-Experimental Lakes Area in 2021. TWA concentrations presented relative to average tadpole sampling event exposure days, presented in ng/L.

Exposure Day	Treatment	Enclosure	TWA concentration at Sampling Time (ng/L)
7	Not oiled	E2	29.11
7	Not oiled	E5	23.23
7	Oiled	E1	1130.42
7	Oiled	E3	922.27
7	Oiled	E4	809.67
13	Not oiled	E2	24.32
13	Not oiled	E5	21.04
13	Oiled	E1	861.39
13	Oiled	E3	806.77
13	Oiled	E4	689.06

Table 2-3 Survivorship of tadpoles housed in four cages within the *in-situ* enclosures. Data are expressed as mean and standard deviation.

Treatment	Total Addition ± SD	Known Mortality ± SD	Unknown Fate ± SD	Enclosure Volume (L) ± SD	Tadpole Density (Tadpoles/L)	Survival (%) ± SD
Not Oiled	57.3 ± 0.6	3.0 ± 0.0	4.0 ± 0.0	29,084.5 ± 1,333.1	0.9	87.8 ± 4.7
Oiled	59.0 ± 1.0	3.7 ± 4.6	6.3 ± 0.3	38,045.9 ± 1,065.9	0.9	83.1 ± 6.1

Table 2-4 Mean apical endpoints measured (Gosner stage, snout-vent length, total length, and weight) within each *in-situ* enclosure at the IISD-Experimental Lakes Area from June 18, 2021-July 23,2021, partitioned by sampling event. Data are expressed as mean and standard deviation.

Sampling Event (GS Class)	Treatment	Gosner Stage ± SD	Snout-Vent Length ± SD (mm)	Total Length ± SD (mm)	Weight ± SD(g)
First event (GS 34-36)	Not oiled	36.2 ± 1.3	18.9 ± 1.1	53.0 ± 3.6	1.5 ± 0.2
First event (GS 34-36)	Oiled	36.8 ± 1.5	20.0 ± 1.4	55.6 ± 3.5	1.6 ± 0.3
Second event (GS 40-42)	Not oiled	40.7 ± 0.7	21.3 ± 1.9	60.5 ± 3.3	1.9 ± 0.3
Second event (GS 40-42)	Oiled	40.6 ± 0.9	21.3 ± 2.1	59.9 ± 3.8	2.0 ± 0.3

Table 2-5 Mean bioaccumulation of total, parent and alkylated polycyclic aromatic compound contents (ng/gram of lipid) amongst tadpole tissues. Data from the treatments sampled in the *in-situ* enclosures at the IISD-Experimental Lakes Area from June 18, 2021- July 23,2021, partitioned by sampling event. Data are expressed as mean and standard deviation. Concentrations are corrected for lipid content.

Sampling Event (GS Class)	PAC Group	Treatment	Mean Concentration (Nanogram/ Gram Lipid) ± SD
First event (GS 34–36)	Total PACs	Not oiled	1.03 ± 0.65
First event (GS 34–36)	Total PACs	Oiled	5.72 ± 1.90
Second event (GS 40–42)	Total PACs	Not oiled	0.26 ± 0.23
Second event (GS 40–42)	Total PACs	Oiled	2.36 ± 1.50
First event (GS 34–36)	Parent PACs	Not oiled	0.37 ± 0.35
First event (GS 34–36)	Parent PACs	Oiled	0.61 ± 0.22
Second event (GS 40–42)	Parent PACs	Not oiled	0.05 ± 0.04
Second event (GS 40–42)	Parent PACs	Oiled	0.22 ± 0.12
First event (GS 34–36)	Alkylated PACs	Not oiled	1.40 ± 0.87
First event (GS 34–36)	Alkylated PACs	Oiled	6.33 ± 2.02
Second event (GS 40–42)	Alkylated PACs	Not oiled	0.32 ± 0.25
Second event (GS 40–42)	Alkylated PACs	Oiled	2.58 ± 1.61

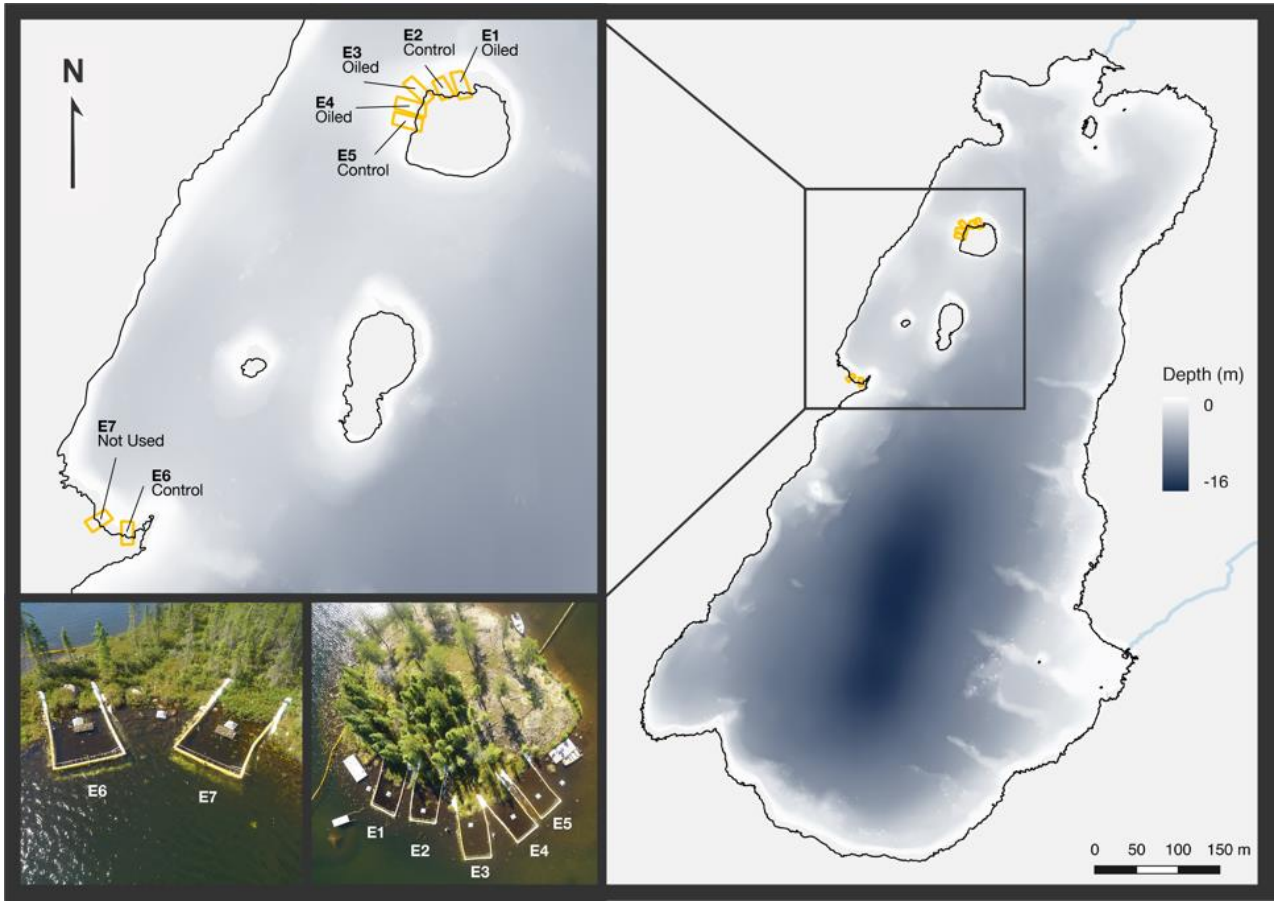


Figure 2-1 Top Right/ Left: Bathymetric map of Lake 260 at the IISD – Experimental Lakes Area. Contour lines represent one meter depth intervals. a 32.76-hectare freshwater research lake, located in northwestern Ontario at the IISD-ELA (49°41'44.4"N 93°46'04.1"W). Bottom Left: Aerial image of enclosures along the shoreline of L260, with assigned treatments. Each pictured enclosure reached a max depth of 1.5-2 meters and were in peat/organic substrate. Total volumes were greater than 20,000 litres of lake water. Blue shading (E2, E5, E6 represent reference, and not oiled enclosures, grey shading on E1, E3, E4 represents treatment/ oiled enclosures, Lake reference was outside of enclosures and was not oiled. Image credit Dr. Scott Higgins, 2021.

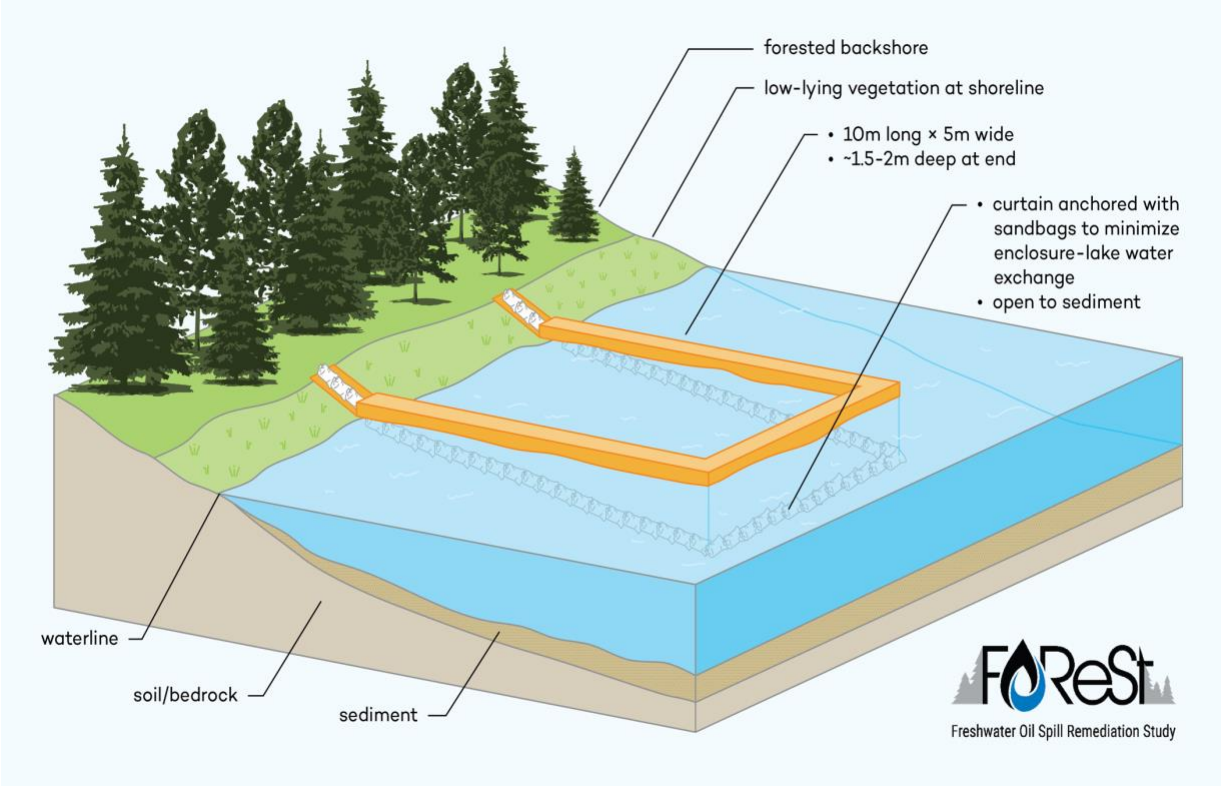


Figure 2-2 Schematic cross-sectional view of the enclosures along the shoreline of Lake 260 at the IISD – Experimental Lakes Area. Enclosure reached a max depth of 1.5-2 meters and were in peat/organic substrate. Total volumes were greater than 20,000 litres of lake water (Palace et al., 2021).

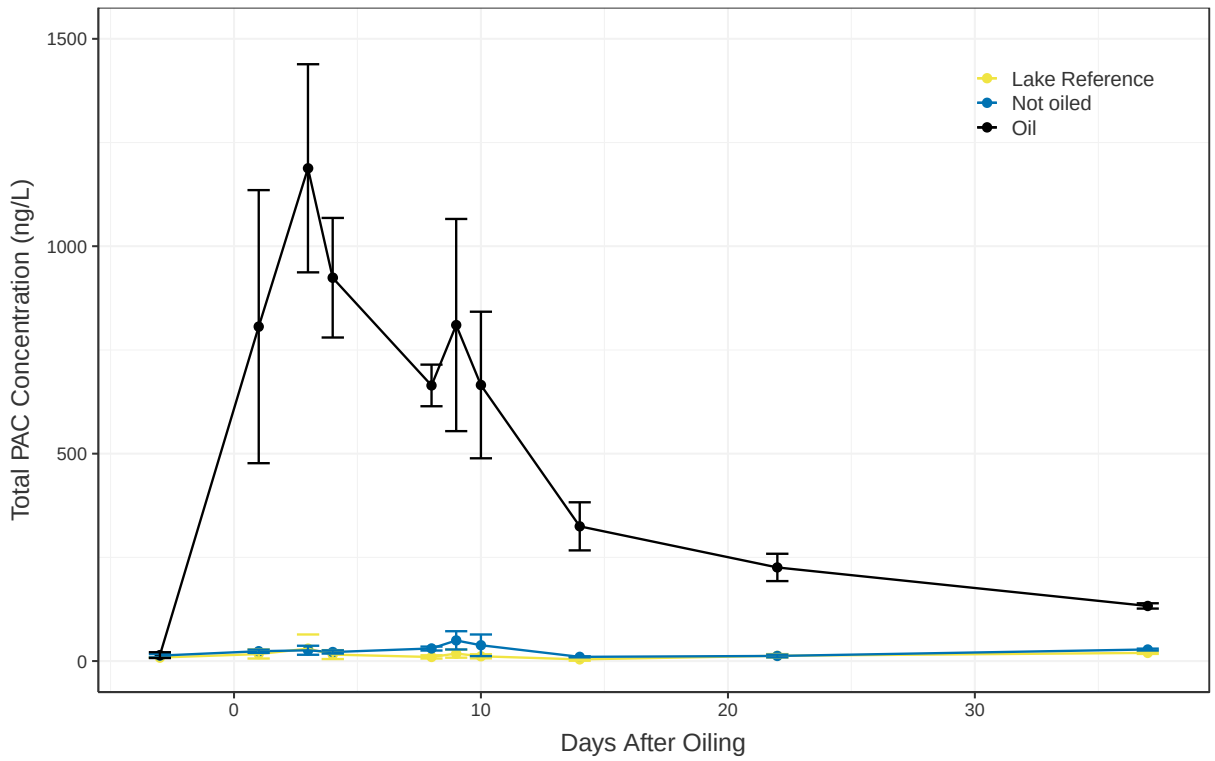


Figure 2-3 Total polycyclic aromatic compound (tPAC) concentrations (in ng/L) of treated 36 hour weathered conventional heavy crude oil (CHV) applied to treatment enclosures at Lake 260 at the IISD-Experimental Lakes Area on June 25, 2021. Plot represented change in tPAC concentrations over duration of tadpole exposure period within the *in-situ* enclosures. data are expressed as mean and standard deviation.

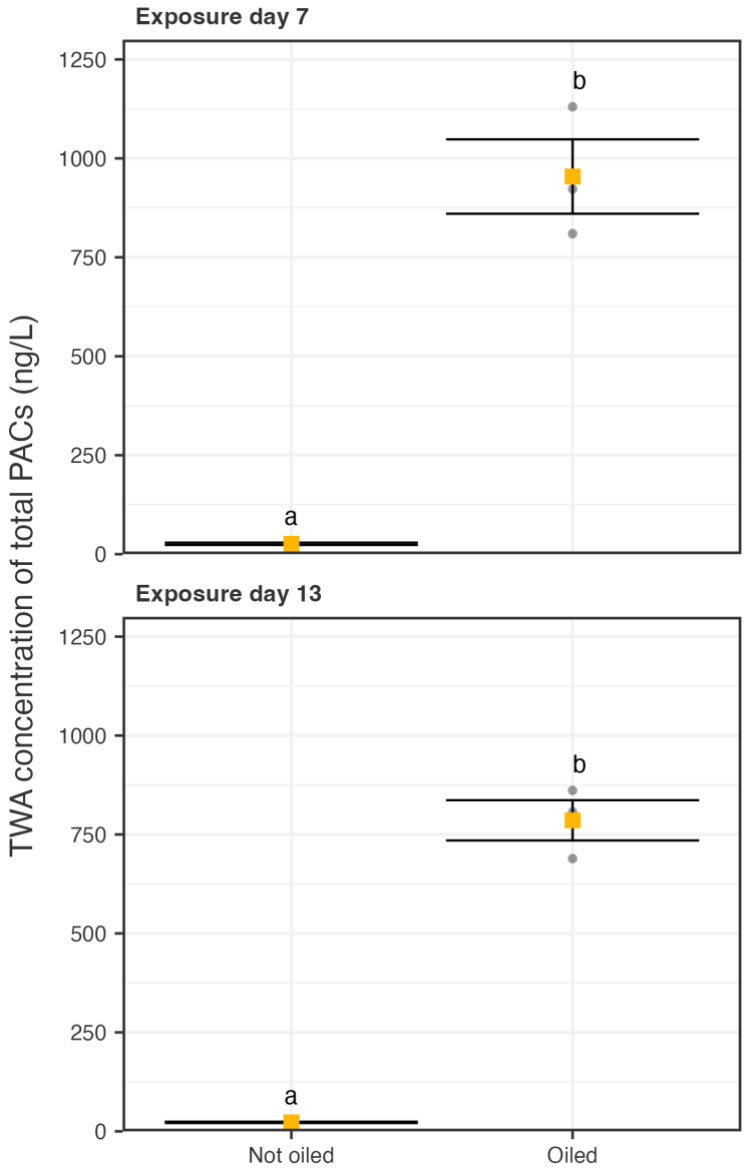


Figure 2-4 Time weighted average (TWA) concentrations of total polycyclic aromatic compounds (tPACs) from CHV oil sampled from enclosure water at the IISD-Experimental Lakes Area in 2021. TWA concentrations presented relative to average tadpole sampling event exposure days, presented in ng/L. Letter represent statistical significance. The multiple comparisons test data for both exposure day indicate significant differences in TWA concentrations of tPACs across the treatments, not oiled (reference) and oiled treatments across both exposure days.

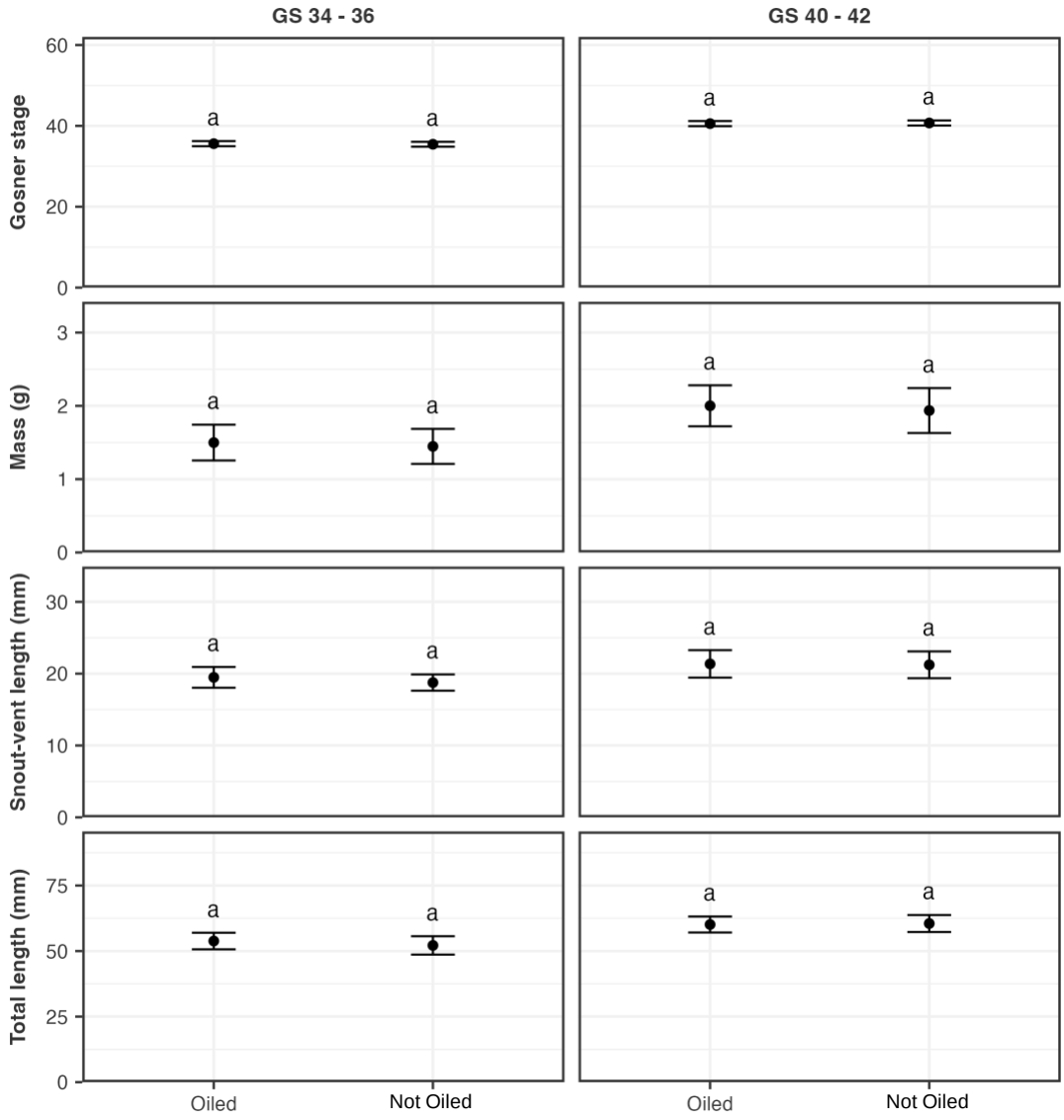


Figure 2-5 Multiple Comparison test plot on apical endpoints measured (Gosner stage, snout-vent length, total length, and weight) within each *in-situ* enclosure at the IISD-Experimental Lakes Area from June 18, 2021- July 23, 2021, partitioned by sampling event. Multiple comparisons test indicates that for tadpoles in GS class 34-36 and GS class 40-42, there were no significant differences between the not oiled (reference) and oiled groups (treatment) across all measured apical endpoints. Letters represent statistical significance; data are expressed as mean and standard deviation.

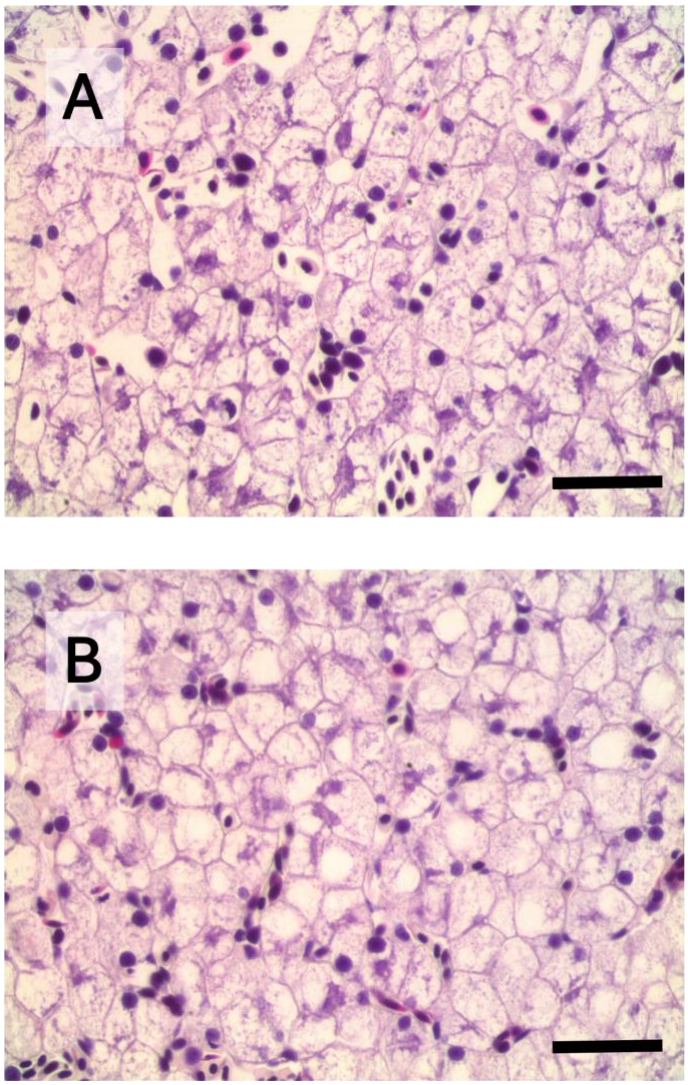


Figure 2-6 Photomicrographs of livers from wood frog (*L. sylvaticus*) tadpoles exposed to various treatments, within each *in-situ* enclosure at the IISD-Experimental Lakes Area from June 18, 2021-July 23, 2021. (A) Not oiled, (B) Oiled. Scale bar = 50 microns.

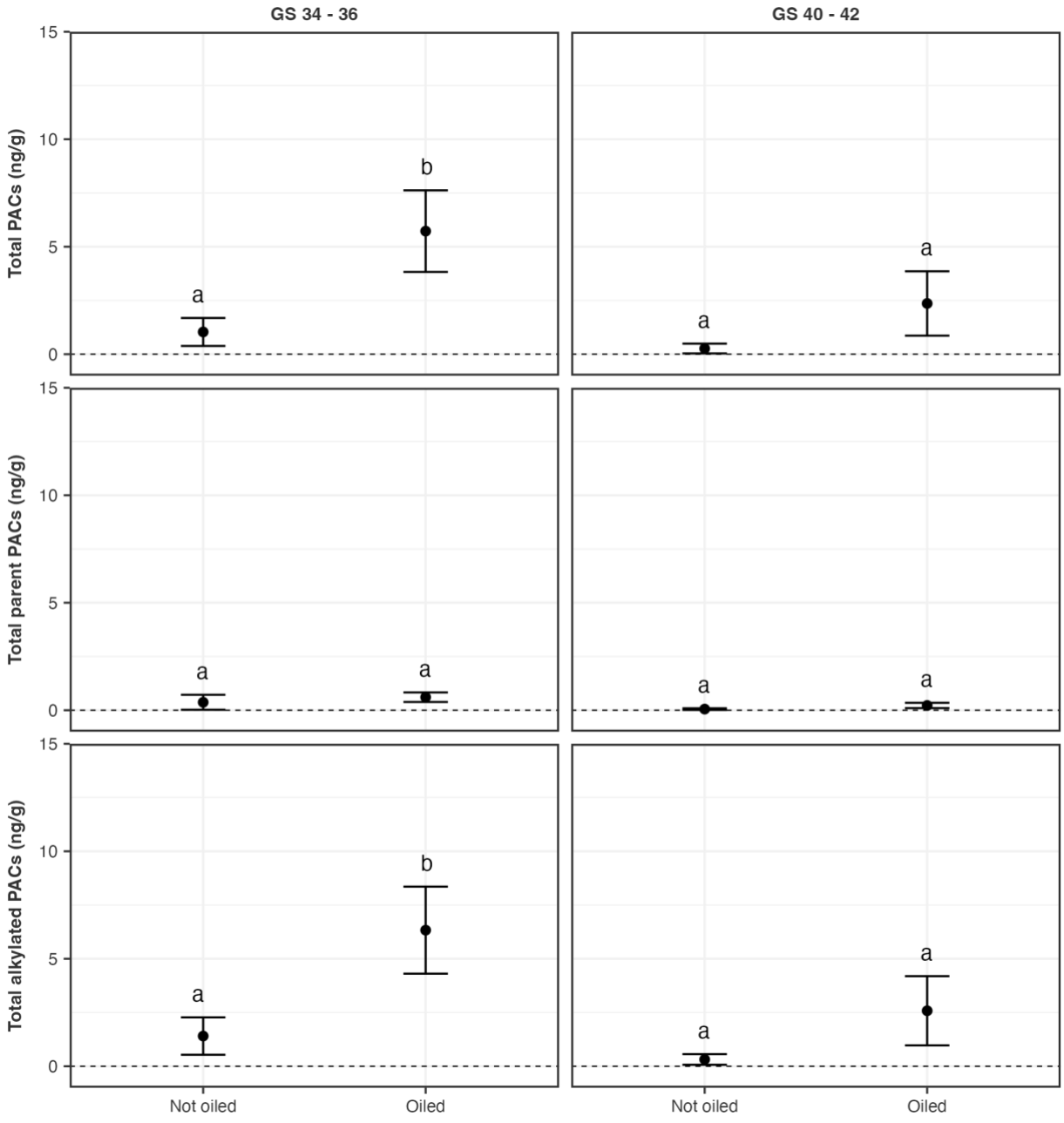


Figure 2-7 Multiple Comparisons figure of tPAC data in tadpole tissues partitioned by GS Class. Total, parent and alkylated polycyclic aromatic compound contents (ng/gram of lipid) amongst tadpoles from the three treatments. Statistical differences were noted in the first sampling event for oiled compared to not oiled treatments, for both tPACs and total alkylated PACs. Concentrations are corrected for lipid content; letters represent statistical significance and tPAC mean and SD presented.

SI Table 2-1 Volumes of the in-situ enclosures along the shoreline of Lake 260 at the IISD – Experimental Lakes Area. Data are expressed as mean and standard deviation.

Enclosure	Intercept	Intercept Slope	Slope	P-Value Slope	RSquare	Initial Enclosure Volume ± SD (L)
E1	561.84	109.01	-3.53 ± 48.76	5.15	0	32,005 ± 6,209
E2	677.49	113.39	-19.54 ± 50.62	5.97	0.04	26,542 ± 4,442
E3	385.85	53.96	3.76 ± 24.03	7.15	0.01	46,603 ± 6,517
E4	506.13	64.63	7.93 ± 28.72	7.83	0.02	35,528 ± 4,536
E5	581.76	60.02	-9.48 ± 35.26	9.69	0.02	30,909 ± 3,188
E6	603.39	35.99	20.34 ± 21.08	16.77	0.24	29,801 ± 1,778

SI Table 2-2 DO (mg/L), DO (%), pH, Specific Conductivity (mS/cm), Temperature (°C) within the (n=7) rearing tubs at the IISD-Experimental Lakes Area from May 4th - June 18th, 2021. Parameters were measured daily with a YSI handheld probe. Data are expressed as mean and standard deviation.

Parameter	**Mean Parameter ± SD**
DO (mg/L)	10.5 ± 6.0
DO (%)	92.9 ± 16.4
pH	6.3 ± 0.3
Specific Conductivity (mS/cm)	57.1 ± 48.5
Temperature (°C)	12.5 ± 5.2

SI Table 2-3 List of Polycyclic aromatic compounds measured for analysis

Compound	
1,7-Dimethylphenanthrene	C1 Pyrene
1,8-Dimethylphenanthrene	C2 Dibenzothiophene
1-Methylnaphthalene	C2 Fluorene
1-Methylphenanthrene	C2 Naphthalene
2,6-Dimethylphenanthrene	C2 Phenanthrene
2-Methylnaphthalene	C2 Pyrene
2-Methylphenanthrene	C3 Chrysene
3,6-Dimethylphenanthrene	C3 Dibenzothiophene
3-Methylphenanthrene	C3 Naphthalene (Trimethyl)
9/4-Methylphenanthrene	C3 Phenanthrene
Acenapthene	C4 Naphthalene (Tetramethyl)
Acenapthylene	C4 Phenanthrene
Anthracene	Chrysene
Benz[a]anthracene	Dibenzo[a,h]anthracene
Benzo[a]pyrene	Dibenzothiophene
Benzo[b]fluoranthene	Fluoranthene
Benzo[g,h,i]perylene	Fluorene
Benzo[k]fluoranthene	Indeno[1,2,3-c,d]pyrene
C1 Benzo[a]pyrene	Naphthalene
C1 Chrysene	Phenanthrene
C1 Dibenzothiophene	Pyrene
C1 Fluorene	Retene

SI Table 2-4 Mean temperature measured by Hobo loggers within each enclosure at the IISD-Experimental Lakes Area from June 18, 2021- July 23, 2021, partitioned by treatment. Temperatures were recorded with a HOBO temperature logger, every 5 minutes. Data are expressed as mean and standard deviation.

Treatments	Temperature ± SD (°C)
Not oiled	23.3 ± 3.1
Oiled	23.2 ± 2.9

SI Table 2-5 Shapiro-Wilk test for normality on the mean temperatures (°C) within the enclosures at the IISD-Experimental Lakes Area. Temperatures were recorded with a HOBO temperature logger, every 5 minutes from June 18, 2021- July 23, 2021.

Treatment	Statistic	P-Value
Not oiled	0.95	0.57
Oiled	0.89	0.36

SI Table 2-6 F-Test to compare variances on the mean temperatures (°C) within the enclosures at the IISD-Experimental Lakes Area. Temperatures were recorded with a HOBO temperature logger, every 5 minutes from June 18, 2021- July 23,2021.

Estimate	num.df	den.df	Statistic	P-Value	Confidence- Low	Confidence- High
2.56	2	2	2.56	0.56	0.07	99.71

SI Table 2-7 Two Sample Students T-Test on the mean temperatures (°C) within the enclosures at the IISD-Experimental Lakes Area. Temperatures were recorded with a HOBO temperature logger, every 5 minutes from June 18, 2021- July 23,2021.

Estimate	Estimate 1	Estimate 2	Statistic	P-Value	Confidence Low	Confidence High
0.03	23.25	23.22	0.14	0.89	-0.61	0.67

SI Table 2-8 DO (mg/L), DO (%), pH, Specific Conductivity (mS/cm), Temperature (°C) partitioned by enclosure treatment at the IISD-Experimental Lakes Area. Parameters were measured with a YSI handheld probe from June 18, 2021- July 23, 2021. Data are expressed as mean and standard deviation.

Treatment	Dissolved Oxygen ± SD (mg/L)	Dissolved Oxygen ± SD (%)	pH ± SD	Specific Conductivity ± SD (mS/cm)	Temperature ± SD (°C)
Not oiled	6.9 ± 0.7	74.9 ± 6.5	6.3 ± 0.3	29.1 ± 3.0	19.7 ± 1.6
Oiled	6.6 ± 1.1	73.0 ± 12.4	6.4 ± 0.3	33.1 ± 7.8	20.3 ± 1.9

SI Table 2-9 Shapiro-Wilk test for normality for DO (mg/L), DO (%), pH, Specific Conductivity (mS/cm), Temperature (°C) partitioned by enclosure treatment at the ISD-Experimental Lakes Area. Parameters were measured with a YSI handheld probe from June 18, 2021- July 23, 2021.

Parameter	Treatment	Statistic	P-Value
Dissolved Oxygen (%)	Not oiled	0.95	0.58
Dissolved Oxygen (%)	Oiled	0.98	0.73
Dissolved Oxygen (mg/L)	Not oiled	1	0.90
Dissolved Oxygen (mg/L)	Oiled	0.96	0.62
Specific Conductivity (mS/cm)	Not oiled	0.96	0.62
Specific Conductivity (mS/cm)	Oiled	0.98	0.72
Temperature (°C)	Not oiled	0.97	0.68
Temperature (°C)	Oiled	0.96	0.61
pH	Not oiled	0.99	0.84
pH	Oiled	0.75	0.00

SI Table 2-10 F-Test to compare two variances for DO (mg/L), DO (%), pH, Specific Conductivity (mS/cm), Temperature (°C) partitioned by enclosure treatment at the ISD-Experimental Lakes Area. Parameters were measured with a YSI handheld probe from June 18, 2021- July 23, 2021

Parameter	Estimate	Number DF	Den DF	Statistic	P-Value	Confidence Low	Confidence High
Dissolved Oxygen (mg/L)	0.09	2	2	0.09	0.16	0	3.37
Dissolved Oxygen (%)	0.1	2	2	0.1	0.19	0	4.02
pH	37	2	2	37	0.05	0.95	1443
Specific Conductivity (mS/cm)	0.87	2	2	0.87	0.93	0.02	33.98
Temperature (°C)	1.48	2	2	1.48	0.81	0.04	57.81

SI Table 2-11 Two Sample Students T-Test for DO (mg/L), DO (%), pH, Specific Conductivity (mS/cm), Temperature (°C) partitioned by enclosure treatment at the ISD-Experimental Lakes Area. Parameters were measured with a YSI handheld probe from June 18, 2021- July 23, 2021

Parameter	Estimate	Estimate 1	Estimate 2	Statistic	P-Value	Confidence Low	Confidence High
Dissolved Oxygen (mg/L)	0.34	7.43	7.09	0.76	0.49	-0.91	1.6
Dissolved Oxygen (%)	3.68	81.16	77.48	0.71	0.52	-10.72	18.08
pH	-0.02	6.21	6.24	-1.14	0.32	-0.08	0.03
Specific Conductivity (mS/cm)	-3.91	33.93	37.83	-0.87	0.43	-16.35	8.54
Temperature (°C)	-0.19	19.73	19.92	-0.66	0.55	-0.99	0.61

SI Table 2-12 Water nutrient parameters within the in-situ enclosures. Nutrient parameter conditions relevant to tadpoles include alkalinity (Alk), ammonia (NH3), calcium (Ca), chlorine (Cl2), dissolved inorganic carbon (DIC), dissolved organic carbon (DOC), particulate carbon (PC), particulate phosphorus (PP), particulate nitrogen (PN), total dissolved phosphorus TDP), total dissolved nitrogen (TDN), total dissolved solids (TDS), total suspended solids (TSS). Water nutrient parameters were sampled four times between June 18, 2021- July 23, 2021, and analyzed by the IISD-ELA Chemistry Laboratory at the IISD-Experimental Lakes Area. Data are expressed as mean and standard deviation.

Parameter	Treatment	Mean Parameter ± SD
Alkalinity (Alk)	Not oiled	208.35 µEq/L ± 44.51
Alkalinity (Alk)	Oiled	252.69 µEq/L ± 54.27
Ammonia (NH3)	Not oiled	9.19 µg/L ± 3.03
Ammonia (NH3)	Oiled	21.05 µg/L ± 26.17
Calcium (Ca)	Not oiled	3.03 mg/L ± 0.57
Calcium (Ca)	Oiled	3.73 mg/L ± 0.74
Chlorine (Cl2)	Not oiled	0.23 mg/L ± 0.06
Chlorine (Cl2)	Oiled	0.19 mg/L ± 0.08
Dissolved Inorganic Carbon (DIC)	Not oiled	263.33 µmol/L ± 59.05
Dissolved Inorganic Carbon (DIC)	Oiled	330.38 µmol/L ± 72.61
Dissolved Organic Carbon (DOC)	Not oiled	671.52 µmol/L ± 65.68
Dissolved Organic Carbon (DOC)	Oiled	730.00 µmol/L ± 126.84
Particulate Carbon (PC)	Not oiled	639.81 µg/L ± 382.29
Particulate Carbon (PC)	Oiled	594.81 µg/L ± 255.99
Particulate Nitrogen (PN)	Not oiled	71.13 µg/L ± 37.35
Particulate Nitrogen (PN)	Oiled	63.14 µg/L ± 23.87
Particulate Phosphorus (PP)	Not oiled	5.19 µg/L ± 2.14
Particulate Phosphorus (PP)	Oiled	5.14 µg/L ± 2.08
Total Dissolved Nitrogen (TDN)	Not oiled	406.91 µg/L ± 29.33
Total Dissolved Nitrogen (TDN)	Oiled	438.095 µg/L ± 73.71
Total Dissolved Phosphorus (TDP)	Not oiled	5.371 µg/L ± 0.96
Total Dissolved Phosphorus (TDP)	Oiled	5.933 µg/L ± 2.18
Total Dissolved Solids (TDS)	Not oiled	59.381 mg/L ± 90.20
Total Dissolved Solids (TDS)	Oiled	39.524 mg/L ± 27.02
Total Suspended Solids (TSS)	Not oiled	1.614 mg/L ± 1.28
Total Suspended Solids (TSS)	Oiled	1.486 mg/L ± 0.88

SI Table 2-13 Shapiro-Wilk test for normality for water nutrient parameters within the *in-situ* enclosures. Nutrient parameter conditions relevant to tadpoles include alkalinity (Alk), ammonia (NH3), calcium (Ca), chlorine (Cl2), dissolved inorganic carbon (DIC), dissolved organic carbon (DOC), particulate carbon (PC), particulate phosphorus (PP), particulate nitrogen (PN), total dissolved phosphorus TDP), total dissolved nitrogen (TDN), total dissolved solids (TDS), total suspended solids (TSS). Water nutrient parameters were sampled four times between June 18, 2021- July 23, 2021, and analyzed by the IISD-ELA Chemistry Laboratory at the IISD-Experimental Lakes Area.

Parameter	Treatment	Statistic	P-Value
Alkalinity (Alk)	Not oiled	0.97	0.66
Alkalinity (Alk)	Oiled	0.9	0.4
Ammonia (NH3)	Not oiled	0.99	0.77
Ammonia (NH3)	Oiled	0.76	0.01
Calcium (Ca)	Not oiled	0.97	0.68
Calcium (Ca)	Oiled	0.93	0.48
Chlorine (Cl2)	Not oiled	0.75	0
Chlorine (Cl2)	Oiled	0.92	0.46
Dissolved Inorganic Carbon (DIC)	Not oiled	1	0.95
Dissolved Inorganic Carbon (DIC)	Oiled	0.82	0.15
Dissolved Organic Carbon (DOC)	Not oiled	0.95	0.58
Dissolved Organic Carbon (DOC)	Oiled	0.9	0.4
Particulate Carbon (PC)	Not oiled	0.94	0.51
Particulate Carbon (PC)	Oiled	1	0.9
Particulate Nitrogen (PN)	Not oiled	1	0.99
Particulate Nitrogen (PN)	Oiled	0.9	0.38
Particulate Phosphorus (PP)	Not oiled	0.92	0.44
Particulate Phosphorus (PP)	Oiled	1	0.99
Total Dissolved Nitrogen (TDN)	Not oiled	1	0.87
Total Dissolved Nitrogen (TDN)	Oiled	0.98	0.75
Total Dissolved Phosphorus (TDP)	Not oiled	1	1
Total Dissolved Phosphorus (TDP)	Oiled	0.94	0.52
Total Dissolved Solids (TDS)	Not oiled	0.77	0.05
Total Dissolved Solids (TDS)	Oiled	0.88	0.31
Total Suspended Solids (TSS)	Not oiled	0.94	0.52
Total Suspended Solids (TSS)	Oiled	0.86	0.27

SI Table 2-14 F-Test to compare two variances on water nutrient parameters within the *in-situ* enclosures. Nutrient parameter conditions relevant to tadpoles include alkalinity (Alk), ammonia (NH3), calcium (Ca), chlorine (Cl2), dissolved inorganic carbon (DIC), dissolved organic carbon (DOC), particulate carbon (PC), particulate phosphorus (PP), particulate nitrogen (PN), total dissolved phosphorus TDP), total dissolved nitrogen (TDN), total dissolved solids (TDS), total suspended solids (TSS). Water nutrient parameters were sampled four times between June 18, 2021- July 23, 2021, and analyzed by the IISD-ELA Chemistry Laboratory at the IISD-Experimental Lakes Area.

Parameter	Estimate	Number DF	Den DF	Statistic	P-Value	Confidence Low	Confidence High
Alkalinity (Alk)	0.59	2	2	0.59	0.74	0.02	22.89
Ammonia (NH3)	0	2	2	0	0	0	0.01
Calcium (Ca)	0.52	2	2	0.52	0.68	0.01	20.1
Chlorine (Cl2)	0.48	2	2	0.48	0.65	0.01	18.75
Dissolved Inorganic Carbon (DIC)	0.82	2	2	0.82	0.9	0.02	31.98
Dissolved Organic Carbon (DOC)	0.12	2	2	0.12	0.21	0	4.63
Particulate Carbon (PC)	1.99	2	2	1.99	0.67	0.05	77.7
Particulate Nitrogen (PN)	1.99	2	2	1.99	0.67	0.05	77.47
Particulate Phosphorus (PP)	3.36	2	2	3.36	0.46	0.09	130.85
Total Dissolved Nitrogen (TSN)	0	2	2	0	0.01	0	0.16
Total Dissolved Phosphorus (TDP)	0	2	2	0	0	0	0.02
Total Dissolved Solids (TDS)	7.15	2	2	7.15	0.25	0.18	278.92
Total Suspended Solids (TSS)	0.03	2	2	0.03	0.06	0	1.21

SI Table 2-15 Two Sample Students T-Test on water nutrient parameters within the *in-situ* enclosures. Nutrient parameter conditions relevant to tadpoles include alkalinity (Alk), ammonia (NH3), calcium (Ca), chlorine (Cl2), dissolved inorganic carbon (DIC), dissolved organic carbon (DOC), particulate carbon (PC), particulate phosphorus (PP), particulate nitrogen (PN), total dissolved phosphorus TDP), total dissolved nitrogen (TDN), total dissolved solids (TDS), total suspended solids (TSS). Water nutrient parameters were sampled four times between June 18, 2021- July 23, 2021, and analyzed by the IISD-ELA Chemistry Laboratory at the IISD-Experimental Lakes Area.

Parameter	Estimate	Estimate 1	Estimate 2	Statistic	P-Value	Confidence Low	Confidence High
Alkalinity (Alk)	-44.34	208.35	252.7	-1.08	0.34	-158.7	70.01
Ammonia (NH3)	-11.86	9.19	21.05	-1.1	0.33	-41.74	18.03
Calcium (Ca)	-0.69	3.03	3.73	-1.28	0.27	-2.2	0.81
Chlorine (Cl2)	0.04	0.23	0.19	1.37	0.24	-0.04	0.12
Dissolved Inorganic Carbon (DIC)	-67.05	263.33	330.38	-1.25	0.28	-216.19	82.09
Dissolved Organic Carbon (DOC)	-58.48	671.52	730	-0.75	0.49	-274.89	157.94
Particulate Carbon (PC)	45.01	639.82	594.81	0.3	0.78	-374.46	464.48
Particulate Nitrogen (PN)	7.99	71.13	63.14	0.53	0.63	-34.07	50.05
Particulate Phosphorus (PP)	0.05	5.19	5.14	0.05	0.97	-2.81	2.91
Total Dissolved Nitrogen (TSN)	-31.19	406.91	438.1	-0.72	0.51	-151.23	88.85
Total Dissolved Phosphorus (TDP)	-0.57	5.37	5.94	-0.49	0.65	-3.77	2.64
Total Dissolved Solids (TDS)	19.86	59.38	39.52	0.81	0.47	-48.61	88.33
Total Suspended Solids (TSS)	0.13	1.61	1.49	0.41	0.7	-0.73	0.99

SI Table 2-16 Mean exposure duration of tadpoles within *in-situ* enclosures partitioned by sampling event from June 18, 2021- July 23, 2021. Data are expressed as mean and standard deviation.

Sampling Event (GS Class)	Mean Sampled Study Day ± SD
First event (GS 34-36)	7.0 ± 0.0
Second event (GS 40-42)	11.3 ± 1.5

SI Table 2-17 Multiple Comparison test on time weighted average (TWA) concentrations of total polycyclic aromatic compounds (tPACs) from CHV sampled from enclosure water at the IISD-Experimental Lakes Area in 2021. The multiple comparisons test data show significant differences in TWA concentrations of tPACs across the treatments, not oiled (reference) and oiled treatments across both exposure days.

Exposure Day	Contrast	Estimate	SE	DF	T.Ratio	P.Value
7	Not oiled - Oiled	-927.95	121.311	3	-7.649	0.005
13	Not oiled - Oiled	-763.06	65.654	3	-11.622	0.001

SI Table 2-18 Multiple Comparison test on apical endpoints measured (Gosner stage, snout-vent length, total length, and weight) within each *in-situ* enclosures at the IISD-Experimental Lakes Area from June 18, 2021- July 23,2021, partitioned by sampling event. Multiple comparisons test indicates that for tadpoles in GS class 34-36 and GS class 40-42, there were no significant differences between the not oiled (reference) and oiled groups (treatment) across all measured apical endpoints.

Sampling Event (GS Class)	Endpoint	Contrast	Estimate	SE	DF	T.Ratio	P.Value
First event (GS 34-36)	Weight (g)	Reference - Oiled	-0.048	0.092	3.439	-0.517	0.637
First event (GS 34-36)	Gosner Stage	Reference - Oiled	-0.119	0.266	3.695	-0.446	0.68
First event (GS 34-36)	Snout Vent Length (mm)	Reference - Oiled	-0.722	0.46	3.26	-1.568	0.208
First event (GS 34-36)	Total Length (mm)	Reference - Oiled	-1.47	1.484	3.732	-0.99	0.382
Second event (GS 40-42)	Weight (g)	Reference - Oiled	-0.074	0.095	3.998	-0.78	0.479
Second event (GS 40-42)	Gosner Stage	Reference - Oiled	0.169	0.149	3.992	1.132	0.321
Second event (GS 40-42)	Snout Vent Length (mm)	Reference - Oiled	-0.166	0.453	3.993	-0.365	0.733
Second event (GS 40-42)	Total Length (mm)	Reference - Oiled	0.191	1.082	3.996	0.176	0.869

SI Table 2-19 Mixed Effect Model on apical endpoints measured (Gosner stage, snout-vent length, total length, and weight) within each enclosure at the IISD-Experimental Lakes Area from June 18, 2021- July 23,2021, partitioned by sampling event. F-statistics indicate that there are differences in variability among the different endpoints measured in each GS class, however, none of the F-statistics suggest strong evidence of significant differences.

Sampling Event (GS Class)	Endpoint	npar	sumsq	meansq	Statistic
First event (GS 34-36)	Weight (g)	1	0.017	0.017	0.291
First event (GS 34-36)	Gosner Stage	1	0.074	0.074	0.209
First event (GS 34-36)	Snout Vent Length (mm)	1	4.468	4.468	2.752
First event (GS 34-36)	Total Length (mm)	1	10.637	10.637	1.025
Second Event (GS 40–42)	Weight (g)	1	0.048	0.048	0.609
Second Event (GS 40–42)	Gosner Stage	1	0.489	0.489	1.283
Second Event (GS 40–42)	Snout Vent Length (mm)	1	0.456	0.456	0.134
Second Event (GS 40–42)	Total Length (mm)	1	0.279	0.279	0.031

SI Table 2-20 Multiple comparisons test on bioaccumulation of total, parent and alkylated polycyclic aromatic compound contents (ng/gram of lipid) amongst tadpole tissues. Data from the treatments sampled in the *in-situ* enclosures at the IISD-Experimental Lakes Area from June 18, 2021- July 23,2021, partitioned by sampling event. Statistical differences were noted in the first sampling event for oiled compared to not oiled treatments, for both tPACs and total alkylated PAHs. Concentrations are corrected for lipid content.

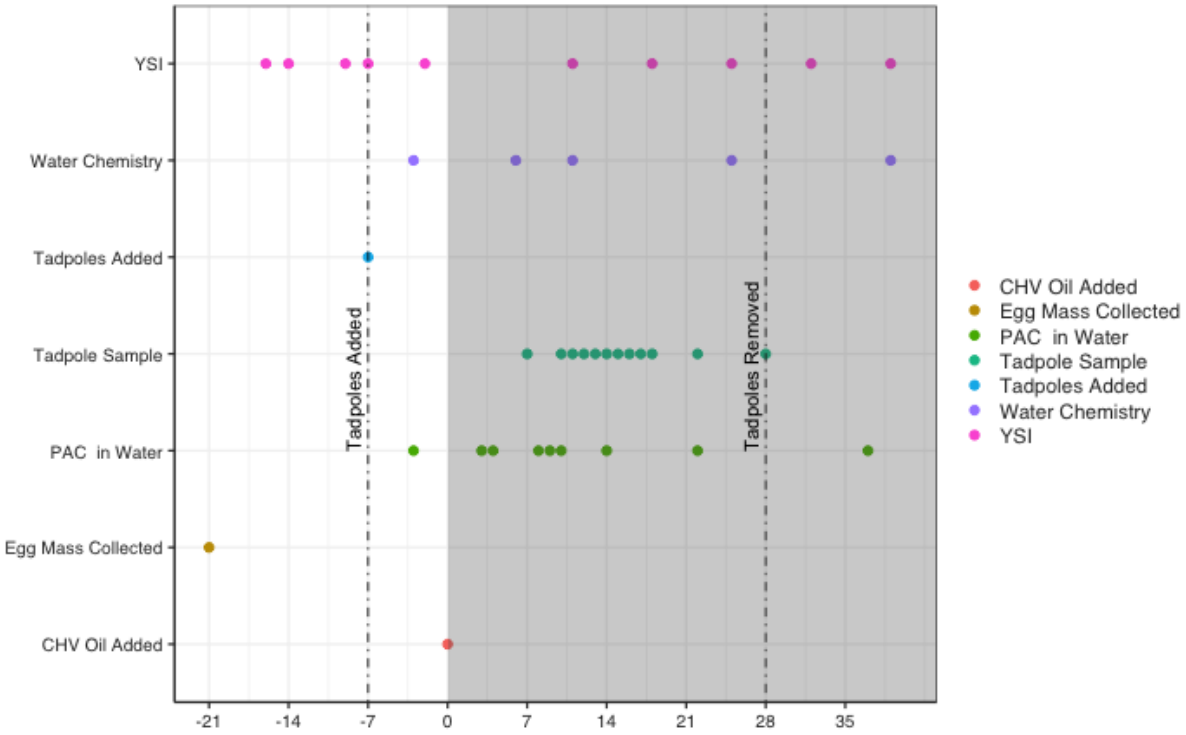
Sampling Event (GS Class)	PAH Group	Contrast	Estimate	SE	DF	T.Ratio	P.Value
First event (GS 34–36)	Alkylated PACs	Not oiled - Oiled	-5.471	1.046	3.974	-5.23	0.006
First event (GS 34–36)	Total PACs	Not oiled - Oiled	-5.148	0.914	3.965	-5.634	0.005
First event (GS 34–36)	Parent PACs	Not oiled - Oiled	-0.296	0.179	3.924	-1.653	0.175
Second event (GS 40–42)	Alkylated PACs	Not oiled - Oiled	-2.265	0.919	4	-2.465	0.069
Second event (GS 40–42)	Total PACs	Not oiled - Oiled	-2.096	0.856	4	-2.449	0.071
Second event (GS 40–42)	Parent PACs	Not oiled - Oiled	-0.169	0.064	4	-2.619	0.059

SI Table 2-21 Mixed Effect Model on bioaccumulation of total, parent and alkylated polycyclic aromatic compound contents (ng/gram of lipid) amongst tadpole tissues. Data from the treatments sampled in the *in-situ* enclosures at the IISD-Experimental Lakes Area from June 18, 2021- July 23,2021, partitioned by sampling event. Concentrations are corrected for lipid content.

Sampling Event (GS Class)	PAH Group	npar	sumsq	meansq	Statistic
First event (GS 34-36)	Alkylated PACs	1	36.774	36.774	27.464
First event (GS 34-36)	Total PACs	1	37.848	37.848	31.916
First event (GS 34-36)	Parent PACs	1	0.18	0.18	2.764
Second event (GS 40–42)	Alkylated PACs	1	2.991	2.991	6.078
Second event (GS 40–42)	Total PACs	1	2.534	2.534	5.998
Second event (GS 40–42)	Parent PACs	1	0.031	0.031	6.857

SI Figure 2-1 A large 600 litre plastic cattle tub filled ¾ with water, and then within each larger tank, three 70 litre tubs, filled ¾ with Lake 260 water. Each smaller 70 Litre tub covered by predator mesh with aeration tubing submerged in the water. Hobo temperature loggers are present in each of the 70 litre tubs and these tubs contained the wood frog egg masses.

SI Figure 2-2 Tadpole cages within *in-situ* enclosures.



SI Figure 2-3 Figure of the sampling timeline for project. Grey represents date of treatment (day 0= June 25th, 2021). The dots represent the frequency of sampling events.

Chapter 3 - Characterizing the Effects of a Conventional Heavy Crude-Derived Water Accommodated Fraction Exposure on the Growth, Development, and Behaviour of Larval Wood Frog (*Lithobates sylvaticus*) exposed in outdoor mesocosms

ABSTRACT

In 2021, a mesocosm study to examine the efficacy of symbiotic plant-microbe relationships for *in-situ* degradation of oil-derived hydrocarbons was conducted using a conventional heavy crude oil (CHV) water-accommodated fraction and engineered floating wetlands (EFWs). This was the Floating Wetland Treatments to Enhance Remediation (FLoWTER) study, conducted at the International Institute for Sustainable Development-Experimental Lakes Area in northwestern Ontario, Canada. As part of the overall FLoWTER study, the research described in this chapter characterized the impacts of oil-spill-impacted waters on the growth, development, histology, and behaviour of wood frog (*Lithobates sylvaticus)* tadpoles, along with an assessment of bioaccumulation. After an exposure time of 28 to 55 study days (July 2nd to July 29th, 2021), there were no statistically significant treatment effects on total length, snout-vent length, and tadpole weight. Bioaccumulation of total polycyclic aromatic compounds (tPAC) in tadpole whole bodies was not statistically different among treatments. Qualitative histology analysis also supports these results, with no apparent differences in hepatic tissues. Behavioural assays investigated the effects of oil presence, with and without plants, on tadpole behaviour throughout the exposure period. Results confirmed the lack of significant treatment effects on total distance, nearest neighbour distance, or percent time in the center on any day of the study, including when a predator cue was introduced. Overall, these findings suggest that the amphibian early life stages studied were not affected by the exposure scenario tested under the conditions and remediation methods applied.

3.1 Introduction

In recent years, Canada has increasingly focused on understanding the potential ecological impacts of oil spills on freshwater ecosystems, prompted by the growing occurrence of such spills and the need for effective remediation strategies (Kheraj, 2020). Industry-standard methods for remediating oil spills, such as applying chemical dispersants, conducting physical and chemical treatments, and using containment and removal techniques are commonly used but can be invasive and may harm sensitive shoreline ecosystems (Whelan et al., 2014; Palace et al., 2021). Recently, there has been a shift toward gentler, less physically disruptive techniques. Bioremediation, which relies on natural processes and microbial activity to degrade oil and other pollutants, is one such approach that aims to reduce environmental disturbance by enabling microbial communities to break down oil compounds gradually (Palace *et al.*, 2021).

The Floating Wetland Treatments to Enhance Remediation (FLoWTER) study built upon the FOReSt study (Chapter 2) to assess the impacts of oil spills in freshwater ecosystems and research the efficacy of engineered floating wetlands in remediation (Stanley, 2024). The FLoWTER study assessed the use of engineered floating wetlands (EFWs) for breaking down CHV-derived hydrocarbons within a CHV water-accommodated fraction (WAF). The EFWs consisted of floating mats that host various aquatic plants whose root systems extend into the water column, enhancing pollutant exposure, uptake and degradation (Rehman *et al.*, 2019). These EFWs can be cost-effective, low-maintenance, and environmentally sustainable for remediation, and they use the natural properties of wetland ecosystems to treat contaminated water, making them suitable for *in-situ* remediation in freshwater systems (Rehman *et al.*, 2018). This approach is a promising solution for oil spill remediation (although minimally researched), as it could offer a less invasive alternative that minimizes habitat disturbance while facilitating the degradation of oil-derived hydrocarbons (Stanley, 2024).

As part of the overall FLoWTER study, this chapter has the objective of characterizing the effects of chronic exposure to CHV-derived hydrocarbons (WAF) on the growth and development, histopathology, behaviour, and bioaccumulation in *Lithobates sylvaticus*. It was hypothesized that exposure to oil-derived hydrocarbons could have an impact on developmental parameters, time to metamorphosis, and tadpole activity, potentially affecting long-term survivability if exposed early

in wood frog tadpole's larval stages. This study also investigated the influence of engineered floating wetlands (EFW) presence or absence on mitigating these responses to CHV-derived hydrocarbons, with the null hypothesis positing that there would be no significant differences in developmental parameters, activity, and bioaccumulation of CHV-derived hydrocarbons in wood frogs resulting from chronic exposure during larval stages. Additionally, behavioural endpoints in amphibians are increasingly recognized as effective measures of pollutant exposure. Subtle changes in behaviour can indicate ecological stress before more severe toxicological effects become apparent. Behavioural metrics such as swimming ability, activity levels, space use, and social behaviour have been linked to exposure effects in amphibians, potentially impacting predator avoidance and foraging efficiency (Peterson *et al.*, 2017; Lee-Jenkins and Robinson, 2018; Bertram *et al.*, 2024).

Through the characterization of toxicological effects on wood frogs exposed to WAF in freshwater environments, this study aimed to contribute to the expanding field of freshwater ecotoxicology by evaluating the impacts of oil spill scenarios and their bioremediation on amphibian health and behaviour. Using the tadpoles within the mesocosms, these insights are intended to inform the development of more effective and ecologically compatible remediation practices, with broader implications for protecting sensitive amphibian populations in freshwater environments in North America.

3.2 Materials and Methods

3.2.1 General Experimental Design

The FLoWTER study commenced in the summer of 2021 and consisted of 26 land-based, outdoor mesocosms with different ratios of aquatic plants within their engineered floating wetlands (EFWs), in addition to treatments of CHV-WAF. The objective of the larger study use these 26 tanks was to develop specific protocols for optimizing the microbiome associated with EFW rhizomes to enhance the degradation of CHV-derived hydrocarbons (WAF) in freshwater shoreline environments.

The mesocosms (also referred to as ‘tanks’) were located on an open area of bedrock at the IISD-ELA (49°42'20.3"N 93°45'50.9" W) and were contained within an 8-foot fence to keep wildlife out of the study area (SI Figure 3-1). The mesocosms consisted of commercially-available stock tanks (product # ARM-10139, Ace-Rotomold, Hospers, Iowa, USA) with a 2.5 m diameter, 0.6 m depth, and a maximum capacity of 2,300 L, constructed USDA-compliant low-density polyethylene (SI Figure 3-2, SI Figure 3-3). The study site was chosen for its proximity to one of the IISD-ELA lakes (Lake 259), which could be used to fill the tanks to maintain water levels throughout the experiment.

Within the overall FLoWTER study, this work focused on a subset of nine tanks along three experimental groups (SI Figure 3-2, SI Figure 3-3). All nine tanks used for this study contained an engineered floating wetland (EFW) consisting of a 1-meter squared matrix of polyethylene fibres, with foam for added buoyancy, and coated in polyurea. The structure were outfitted with several holes to allow for the planting and growth of wetland vegetation in contact with the water column. The EFWs within the tanks were placed in a central location where they remained in place thanks to a weighted anchor line. The EFWs in six of the nine tanks were planted with aquatic macrophytes between 2019 and 2020 to allow for an established community consisting of varying proportions of cattails and two different sage species (SI Figure 3-2, SI Figure 3-3). The remaining three tanks contained unplanted engineered floating wetlands (SI Figure 3-2, SI Figure 3-3, SI Table 3-5).

Regarding oil treatments, three of the tanks with planted EFW and the three tanks with the unplanted EFWs received an oil treatment in the form of a conventional crude oil water accommodated fraction (CHV-WAF, see section 3.2.3.2 for more details) each, while the final set of three tanks with planted EFWs remained un-oiled. The addition of the CHV-WAF to the appropriate treatment mesocosms occurred on June 2nd, 2021 (Study Day 0). This way, the three treatments tested in this study can be summarised as 1) oil with plants, 2) oil without plants, and 3) no oil (with plants).

The initial water volumes in the tanks averaged 1,608 L ± SD 57 litres. Mean volumes for the duration of the experiment can be found in (SI Table 3-5). Tadpoles were added to all 26 tanks on May 31st, 2021 (2 days before oil addition), to maintain a consistent grazing pressure in all

tanks and minimize influence in other components of the overall study, which could be affected by changes in this grazing pressure (e.g. biofilm biomass). As previously noted, only data from the subset of the nine mesocosms previously described were used and presented in this thesis.

3.2.2 Wood Frog Rearing and Handling

All work with *Lithobates sylvaticus* (wood frog) was performed under an approved animal care protocol (University of Manitoba Animal Care Protocol FLoWTER F20-007- AC11577). Wood frog egg masses (n=7) were collected from Lake 227 on May 4th, 2021, and transported back to a rearing area (SI Figure 2-1). This setup consisted of two 600-litre plastic cattle tanks that were filled ¾ with water from Lake 240, and then within each larger tank, up to seven separate 70-litre tanks, also filled ¾ with water from Lake 260, held the separate wood frog egg masses to prevent overcrowding. These animals were used for two simultaneously conducted studies (see this chapter and Chapter 2).

Temperature and basic water quality parameters (specific conductivity, pH, alkalinity, and dissolved oxygen levels) were monitored daily in rearing tanks containing the wood frog egg masses. Water in the tanks was gently replaced with fresh water from Lake 260 at least every third day following recommended rearing procedures (Bridges, 1997). Water temperature constantly measured with HOBO loggers during the rearing period ranged from (2.7°C to 30.2°C, mean of 13.7°C ± SD 4.9°C), while outside air temperature fluctuated from (-1.6°C to 34.8°C, mean of 14.4°C ± SD 7.6°C) (refer to Chapter 2.2). The water quality data determined that the rearing tank conditions were acceptable for rearing wood frog tadpoles following OECD guidelines (OECD, 2009, 2015).

Each egg mass was contained in a small hairnet to keep them afloat and self-contained. This allowed for the individual egg masses to be monitored daily for 11 days during the egg stage. The date of which the majority of tadpoles hatched from egg mass occurred on May 15th, 2021, ± SD 3.5 days (determined by the embryo having emerged from the egg mass). Recently hatched tadpoles were fed frozen kale flakes during the pre-exposure holding time.

On May 30th, 2021, before moving the hatched tadpoles to the 9 tanks, a subsample (n=20) of tadpoles were euthanized (using a buffered MS-222 solution), apical endpoints were measured,

and then tadpoles were frozen for pre-exposure baseline determination. During this pre-exposure period, the mean Gosner stage at which tadpoles were sampled was 22.9 (± SD 1.6). The pre-exposure measurements for tadpoles reared in the rearing tanks at the IISD-Experimental Lakes Area from May 4th to May 30th, 2021 (26 days) were as follows: the mean snout-vent length was 5.4 mm (± SD 1.7), mean total length was 14.1 mm (± SD 3.5), and the mean weight was 0.04 g (± SD 0.03).

Any unhatched eggs were then removed from the rearing tanks. For timing feasibility, within the rearing period, all unhatched eggs and larvae mortalities experienced did not have parameters beyond generalized Gosner stage measured. Any hatched tadpoles exhibiting abnormal behaviour or visible malformations were removed and euthanized. Malformations and normal and abnormal behaviour were defined by the *FETAX Assay for Evaluation of Developmental Toxicity* in combination with *The Larval Amphibian Growth and Development Assay* (Mouche *et al.*, 2011; OECD, 2015).

Two days before CHV-WAF addition, tadpoles were transported to the mesocosms in plastic bags (10 cm x 30 cm) containing water from their rearing tanks (originally from L260, the study lake) with a headspace of air (50% volume). Bags were left for roughly 30 minutes in the mesocosms to acclimate tadpoles to water temperature prior to release into the tanks. Eighty (80) early-life stage wood frog tadpoles (27 days old; collected May 4th, Gosner stage 22.9 ± SD 1.6) were stocked into each of the 26 mesocosms on May 31st, 2021.

3.2.3 Water Chemistry

3.2.3.1 General Water Quality

Water within the mesocosms was sampled throughout the experiment and analyzed by the IISD-ELA Chemistry Laboratory on site, following standard procedures similarly to Chapter 2 (Havens *et al.*, 2024). Mesocosms were sampled ten times throughout the tadpole exposure period to assess water quality; specifically, alkalinity (Alk), ammonia (NH3), calcium (Ca), chlorine (Cl2), dissolved inorganic carbon (DIC), dissolved organic carbon (DOC), particulate carbon (PC), particulate phosphorus (PP), particulate nitrogen (PN), total dissolved phosphorus TDP), total

dissolved nitrogen (TDN), total dissolved solids (TDS), total suspended solids. Refer to (SI Figure 3-5) which details which study days sampling of water parameters occurred.

Conditions in the mesocosms were also measured using a Yellow Springs Instruments multiprobe (YSI Professional Series Plus, YSI Quatro, ISD-ISD-DO-COND, 18C100206, Xylem, Yellow Springs, Ohio) for a total of ten times throughout the exposure period (SI Figure 3-5). Within the mesocosms, four HOBO Pendant MX Temperature and Light Data Logger (*Onset Computer Corporation*) were suspended off the engineered floating wetlands ~15 cm beneath the water surface before the addition of CHV-WAF and before study day 0. These temperature loggers within the mesocosms tracked units at 5-minute intervals for the duration of the exposure period (SI Figure 3-5).

3.2.3.2 Creation and Characterization of the Water Accommodate Fraction

A total of 13.3 kg of conventional heavy crude oil (CHV) (provided by the Canadian Association of Petroleum Producers [CAPP] pipeline stock feed) was weathered for 36 hours by pouring the oil into stainless steel holding pans (1.1 m diameter) containing 225L of water from Lake 260 (Kharey *et al.*, 2024). Once weathered, the floating CHV oil was collected for a separate study also being conducted at the International Institute for Sustainable Development Experimental Lakes Area (IISD-ELA) during this period (see Chapter 2). The residual underlying CHV-exposed lake water in the stainless-steel pans was collected via a valve at the bottom of the weathering pan. This water would be the CHV water-accommodated fraction (CHV-WAF), which was collected in stainless-steel containers and transported to the mesocosm installation for use as the treatment solution in the mesocosms. Approximately 16 litres of this CHV-WAF were added into the appropriate treatment tanks on June 2nd, 2021 (Study Day 0)(concentrations found on Table 3-1).

Sampling for total polycyclic aromatic compounds (tPACs) quantification in the mesocosm water was conducted at ten different time points during the tadpole exposure period (SI Figure 3-5) to characterize the exposure in the systems. This sampling schedule aimed to capture potential variations in PAC concentrations throughout the exposure period and assess any changes in tPAC levels in response to the experimental conditions.

The chemistry analysis of tPACs in mesocosm water samples followed the methods developed by (Idowu *et al.*, 2018). Sample preparation and extraction were completed at the IISD-ELA during the 2021 field season, and the samples were then transported to the Center for Oil and Gas Research and Development Laboratory (COGRAD) at the University of Manitoba, where PAC analysis was completed. The methods applied in the analysis allowed for the detection and quantification of numerous PACs, including polycyclic aromatic hydrocarbons (PAHs) and alkyl-PAHs (APAHs) (Idowu *et al.*, 2018). For methods on the study site and further oil chemistry methodology, please also refer to (Kharey *et al.*, 2024). For this study, PAC data will be summarised as the sum of all forty-four analysed PACs, i.e. total PACs (tPAC) (SI Table 3-1).

3.2.4 Tadpole Endpoints

Mortalities and tadpoles experiencing abnormalities and visible morphological malformation were removed and euthanized by overdose in buffered MS-222 solution (0.5 g/L MS-222 and 0.2 g/L sodium bicarbonate in lake water) (Robinson *et al.*, 2017). Malformations and normal and abnormal behaviour were simply defined by *FETAX Assay for Evaluation of Developmental Toxicity* (Mouche *et al.*, 2011; OECD, 2015)

Tadpoles were maintained up to the onset of metamorphosis (tail resorption, mouth parts change, Gosner stage 40-42), at which point they were lethally sampled for morphometric data, and whole-body samples were prepared for either bioaccumulation or fixed for histology (Gosner, 1960). Tadpoles within the mesocosms were free swimming and free feeding up to the onset of metamorphosis. Exposure time until the final endpoint stage ranged from 28 to 55 study days after CVH addition (July 2nd to July 29th, 2021) when tadpoles were removed and euthanized once they reached specific sampling time points denoted by Gosner stage (SI Table 2-16). Target endpoints are identified in section 1.2.5.2 Apical Endpoints below.

3.2.4.1 Survivorship

Survivorship was determined at the end of the exposure period by calculating the total number of tadpoles initially added to the tanks and subsequently adjusting this figure by subtracting the tadpoles used for specific endpoints and pre-determined sampling events.

Additionally, all documented mortalities were recorded throughout the study period. Given the nature of the field experiment, a category for unknown or unaccounted tadpoles was included to address instances where the fate of certain individuals was undetermined. The survivorship rate was then calculated using the formula:

$$Survivorship\ Rate\ (\%) = \left(\frac{Total\ Additions - \left(Known\ Mortalities + Unknown\ Fate\ (Total\ Samples - Known\ Mortalities)\right)}{Total\ Additions}\right) x\ 100$$

3.2.4.2 Apical Endpoints

The classification of the Gosner stage classes in this study followed the methods described and applied by Navarro-Martín *et al.* (2012) and was based on the original Gosner staging system developed by (Gosner, 1960). Apical endpoints (total length, snout-vent length, wet weight and Gosner stage) in tadpoles were measured at three sampling events throughout the study. These developmental categories provided useful time points for sampling that represented key stages in tadpole development, and the phases of the study in which sampling would occur were determined to best align with the developmental stages.

The first sampling event, "pre-metamorphosis" for baseline determination, occurred at GS26, which is denoted by the formation of the operculum (Gosner, 1960; Hersikorn and Smits, 2011; Navarro-Martín *et al.*, 2012). At this point, a subsample of 20 tadpoles were euthanized and sampled. The second stage, "pre-metamorphosis," took place during Gosner stage GS 34-36, when the majority of tadpoles have digit development on their hind limbs, suggesting this is a midpoint stage in wood frog metamorphosis (Gosner, 1960; Reeves *et al.*, 2011; Navarro-Martín *et al.*, 2012). A total of 10 "pre-metamorphic" tadpoles in each mesocosm were collected, euthanized, and sampled. The third measurement and the final exposure sampling time-point (called second-event) occurred just before "metamorphic climax", called Gosner stage 41 when front limb buds (initially the left front limb bud) are just emerging (Taylor and Kollros, 1946; Gosner, 1960; Navarro-Martín *et al.*, 2012). At this stage all remaining tadpoles that reached GS 40-42, tadpoles were collected, euthanized and sampled. In the instance where tadpoles did not reach this final Gosner stage prior to "metamorphic climax", the exposure period was then terminated at 55 days post CHV-WAF addition to the mesocosms.

As tadpoles reached the specified predetermined Gosner Stage, wood frog tadpoles were removed from the mesocosms and placed in one-litre clear plastic containers filled with associated tank water with 25-50% airspace, placed in a cooler and transported back to the IISD-Experimental Lakes Area laboratory. At the laboratory, apical endpoints were recorded after euthanizing by submersion in tricaine methanesulfonate (MS-222; 500 mg/L) in lake water, buffered to a neutral pH = 7 with sodium bicarbonate methanesulfonate (Grant, 2008; Iannella *et al.*, 2017).

The developmental stage was assessed via visible observation of each tadpole and using the staging table for anuran amphibians by Gosner (1960), to the approximate nearest Gosner stage. Once tadpoles were euthanized, individual tadpoles were measured for wet weight, total length to the nearest 1.0 millimetre (distance from nose to end of straightened tail, measured with tadpole on their ventral side) and snout to vent length (measured from the snout to the vent on the tadpole) (Gosner, 1960; OECD, 2009). Wet weight was measured in grams using a Denver Instruments 0.001 g digital scale (Denver Instrument P-403 400g x 0.001g, 115 V External Calibration).

Lengths (total and snout-to-vent) were measured using a digital calliper. Note that tadpoles with tail damage or missing sections, either due to natural causes or that occurred during sampling or transport, were excluded from the total length calculations. This exclusion was necessary because accurately measuring the total length (from snout to tip of tail) would not have been possible in such cases. However, the length from the snout to the vent (snout-vent length) was still recorded for these tadpoles.

3.2.4.3 Histology

A subset of tadpoles from the second sampling endpoint (GS 40-42) were collected for histological assessment. The initial targeted allocation of tadpoles sampled for histology was six tadpoles per mesocosm; however, due to sampling restraints, allocated amounts were not always achieved, evident in the oil with plants mesocosms where only 12 tadpoles were available. In the reference mesocosms, a total of 18 tadpoles were used and lastly, within the oil without plant mesocosms, 19 tadpoles were used (one extra tadpole used in this treatment) (SI Table 3-2).

3.2.4.3.1 Histological Sample Preparation

At the end of each exposure period, sampled tadpoles (GS 40-42) were removed and euthanized and those tadpoles designated for histological analysis were then given a mid-ventral incision through the body wall, placed in individually labelled 50 mL Falcon tubes, and fixed whole in 10% neutral buffered formalin. Later, the formalin fixative was decanted from each sample and replaced with 70% ethanol. The tadpoles were then dissected under a stereoscope to remove the livers, which were placed in labelled histocassettes and stored in 70% ethanol until further processing. The ethanol was decanted and replaced twice to remove excess fixative from the tadpole tissues.

Liver samples were submitted to the University of Manitoba Histology Services Lab for paraffin embedding, and the blocks were sectioned and stained to confirm the methods before proceeding with the full batch of samples. The liver histology samples were dehydrated in graded alcohols, cleared in xylene, and infiltrated with molten paraffin using a Shandon Citadel 2000 Tissue Processor at the University of Manitoba Histology Services Laboratory. The samples were embedded in paraffin blocks in random orientation. The blocks were step-sectioned (5 millilitres thick) at three different levels approximately 100 microns apart, mounted on glass slides, and stained with Harris' hematoxylin and eosin.

3.2.4.3.2 Histological Assessment

Livers were assessed qualitatively for hepatic changes in the experimental treatments relative to the control samples, including changes in hepatocyte size, nuclear abnormalities, staining properties (i.e., the degree of the eosinophilic (pink) or basophilic (blue) staining of the hepatocyte parenchyma), cytoplasmic vacuolation, or any other notable differences. For each tadpole, representative sections were examined across all three levels. Similar methods were employed in previous toxicity studies in fish (Alcaraz *et al.*, 2021a; Alcaraz *et al.*, 2021b; Alcaraz *et al.*, 2022; Colville *et al.*, 2022). Slides were analyzed using a Zeiss Axiostar Plus microscope. Photomicrographs were captured using a Lumenera INFINITY1-1M digital camera, and Teledyne Lumenera Infinity Analyze software (Version 7.0.2.930).

3.2.5 Bioaccumulation of Total Polycyclic Aromatic Compounds

Once measured and euthanized, wood frog tadpoles designated for total polycyclic aromatic compound (tPAC) bioaccumulation analysis were put individually into 50 mL Falcon tubes and immediately placed on dry ice to preserve their integrity and prevent any degradation of the tissues. Samples were transported to the Center for Oil and Gas Research and Development Laboratory (COGRAD) at the University of Manitoba and stored at -80°C. Tadpoles were then prepared following similar methods to those outlined in Bilodeau et al., (2019) and Mundy et al., (2019) and prior to bioaccumulation analysis, the entire gastrointestinal (GI) tract was removed from each tadpole ensuring the quantified PAC concentrations were that of the tadpole body and not from undigested food or sediment within the tadpole GI tract (Bilodeau et al., (2019). To achieve sufficient sample mass after the removal of the GI tract, for the quantification of tPAC bioaccumulation in tadpole tissues, tadpoles from the mesocosms were pooled in groups of four (n=4) from the same treatments and same Gosner stage classes (SI Table 3-3).

3.2.5.1 Tissue Extraction

PAC quantification/extraction within tadpole body tissues followed the protocol of Xia *et al.* 2023 (Xia *et al.*, 2023). Individual tadpoles were extracted using a Precellys Evolution Micro-bead homogenizer. Each batch of 16 samples included a method blank, a freeze-dried *Mytilus edulis* tissue (SigmaAldrich, 2974a) reference material blank and reference material spiked with the recovery internal standard (RIS) for quality assurance and quality control purposes and 13 RIS spiked tadpole samples. The 20 µL of 1 ng/ µL RIS spike solution consisted of deuterated 16 US Environmental Protection Agency (US EPA) priority PACs: d8-naphthalene, d8-acenaphthylene, d10-acenaphthene, d10-fluorene, d10-phenanthrene, d10-pyrene, d12-Benz(a)anthracene, d12-chrysene, d12-benzo(b)fluoranthene, d12-benzo(k)fluoranthene, d12-benzo(a)pyrene, d12-indeno(1,2,3-c,d)pyrene, d14-dibenzo(a,h)anthracene, and d14-benzo(g,h,i)perylene (Balmer *et al.*, 2019).

Spiked samples were then extracted with 8 mL dichloromethane (DCM) using the Precellys microbead homogenizer with 15 mL micro-bead extraction tubes containing 2.5 grams of the 2.8 mm zirconium oxide beads. The homogenate was then centrifuged (Bertin, Thermo Scientific

Sorvall ST16 Refrigerated Centrifuge- Sorvall ST 16/16R) and the bottom DCM layer was transferred to a 60 mL test tube. The extraction was repeated two more times, and the DCM extracts were combined. The microbead homogenizer vial was rinsed twice with hexane, which was added to the final extract pool. Any water from the tissue was removed through the addition of sodium sulfate directly to 60 mL test tube and shaken. The sodium sulfate was removed by filtering it into a sterile 30 mL test tube through glass wool. The sample was reduced to 10 mL using an N-Evap nitrogen evaporator for lipid determination.

3.2.5.2 Lipid Determination and Gel Permeation Chromatography (GPC)

Lipid content was accounted for when measuring and analyzing the bioaccumulation of total polycyclic aromatic compounds. Lipid content is measured as percent lipid within the tadpole tissues and presented as nanogram per gram of tissue (ng/g). This was achieved by pipetting 500 µL of the homogenate sample extract (section 3.2.5.1) onto a pre-weighed aluminum boat; the solvent evaporates, leaving the lipid behind. The weigh boat is weighed again and then the final percent concentration was calculated.

To prepare sample extracts for Gel Permeation Chromatography (GPC), the volumes were reduced to 2.5 mL by nitrogen evaporation, and samples were transferred into a 10 mL test tube with 2x DCM rinses. Final volumes were adjusted to 5 mL. The sample extracts were applied to the GPC column (J2-scientific AccuPrep, Columbia, MO, USA). Column description and GPC run parameters are described in Xia et al. 2023 (Xia *et al.*, 2023). The fraction eluted from the column that contained the analytes of interest were collected in a 250 mL round bottom flask. The clean extract was reduced to 200 µL using a Genevac Rocket evaporator (Ipswich, Suffolk, UK) and N-Evap. The final extract was transferred into a 2 ml amber gas chromatography vial with a 300 µL glass insert and spiked with 20 µL of 1 ng/ µL d-anthracene as the instrument performance internal standard (IPIS). Extracts were stored at 4°C prior to GC analysis.

3.2.5.3 Gas chromatography mass spectrometry/ mass spectrometry for Tadpole Tissues

This was set up specifically for the process of tissues as per the methodology established for biota by Xia *et al.* (2023). The instrument used for the native and d-labelled PAC quantification was an Agilent 7890 GC coupled to a 7000C triple quadrupole mass spectrophotometer fitted with an electron ionization (EI) source. Helium was the carrier gas with sample injection volumes of 1 µL injected by a PAL RSI 85 autosampler. GC/MS/MS conditions and details of the multiple reaction monitoring precursor and product ion transitions for the 16 priority PACs that were analyzed in the tadpole tissue samples can be found in (Xia et al., 2023). Refer to SI Table 3-1 for a list of PACs analysed in this methodology.

3.2.6 Behavioural Response Assessments

The following section describing the methodologies used for the assessment of behavioural alterations in wood frog tadpoles after chemical exposure follows, as closely as possible, the principles described in the EthoCRED framework (Bertram *et al.*, 2024). Unfortunately, the final version of this framework was not available at the time of development of the experimental design or execution of the study, so certain aspects could not be accommodated in the original design. Aspects related to reporting were followed during the writing of this manuscript.

3.2.6.1 Reference Toxicant Behavioural Assessment

To validate the approaches and evaluate the suitability and sensitivity of the methodologies we used for the assessment of behavioural effects in wood frog tadpoles, we conducted a reference-toxicant test using two reference compounds: ethanol (anxiolytic) and caffeine (anxiogenic) (Egan *et al.*, 2009; Sloman and McNeil, 2012). Tadpoles were exposed to eight different concentrations (0.1 – 10 mg/L) plus a control for caffeine and six different concentrations (0.1 – 2.68 % v:v) plus a control for ethanol, each replicated three times, totalling 48 experimental units (SI Table 3-4). Each unit contained five tadpoles (240 tadpoles overall) randomly collected from a pool of the same egg masses used for the mesocosm (Chapter 3) and *in-situ* enclosure studies (Chapter 2). Apical endpoints (total length, snout-vent length, wet weight and Gosner stage) in tadpoles were

measured after the behavioural assays, mean tadpole weight was 0.22g ± SD 0.04g within the caffeine assays, while mean tadpole weight was 0.2g ± SD 0.05g within the ethanol assays.

Exposures were conducted in 1-litre containers for one hour. This design aimed to assess acute pharmacological effects on the tadpoles' behaviour. After 1 hour of exposure, all five tadpoles within a replicate were transferred to a behavioural arena for behavioural assessment. Due to the limited number of arenas available (9), not all experimental units within each test (caffeine or ethanol) could be assessed at once. To allow for the most effective measuring schedule and to account for any variability due to diurnal cycles or other temporal heterogeneity, the three replicates from each treatment were blocked, and the three blocks were staggered temporally with ~30-minute gap between the start of each exposure.

Each behavioural arena consisted of a larger tote (type A Restore Nesting Tote, 40-L) used as an isolation tank inside which a white plastic food-grade container (24 cm x 28 cm x 14 cm with slightly tapered sides) was placed and filled with three litres of exposure media (~ 6 cm of depth) (SI Figure 3-6). These white exposure containers sat on a custom-made wooden frame, which allowed for a ring light to sit under them. Lighting was connected remotely via a 12v car battery and an inverter. Each arena was outfitted with an action camera (Campark X30 native 4K 60fps 20MP Waterproof Video WiFi Action Cameras) attached to the outside of the lid of the isolation tank, with the lens facing a small opening in the lid. The behavioural arenas were constructed in such a way as to provide a consistent backlit image of the tadpoles, ensuring high contrast for digital tracking by the behavioural analysis software. Before observations, the camera setup was arranged to ensure optimal recording conditions and minimize interaction with the subjects during the acclimation and assay periods. At the end of the 1-hour exposure period, the five tadpoles from a particular experimental unit were placed into the white plastic exposure container filled with the same exposure solution to which those tadpoles had been exposed.

During the behavioural trials, tadpoles were left to acclimate in the arena for 10 minutes before the start of recording (Mirza *et al.*, 2013). Recording was conducted for a 10-minute observation period. After recording, tadpoles were returned to their exposure containers. Subsequently, tadpoles were weighed, measured, and photographed from the dorsal view before being euthanized in MS-222. Upon completion of the behavioural assay for one replicate, the next

block of replicates was processed in the same manner. The caffeine and ethanol tests were separately conducted over two consecutive days.

Videos were processed through a behavioural analysis video-tracking software (LoliTrack V. 5, Loligo Systems, Denmark, 2020). Behavioural endpoints quantified included measures of overall activity (total distance travelled), and measures of socialization (distance to nearest neighbour). Additionally, to assess the thigmotaxis behaviour of the tadpoles (their tendency to remain close to walls and limit exposure to open spaces), the arena was digitally divided into two sections, an outer perimeter and an inner "exposed" area representing 50% of the total area of the arena and time spent in each of these areas was measured (Denoël *et al.*, 2012; Carlson and Langkilde, 2013). These values were computed for each tadpole.

3.2.6.2 Behavioural endpoints assessed in the WAF experiment

For the main WAF exposure tank experiment, behavioural observations were conducted weekly on a subset of five tadpoles from each mesocosm. The same behavioural arenas described above were used, with each of the nine arenas being allocated to a specific mesocosm to prevent cross-contamination. Equipment used included aquarium dip nets for capturing and transferring tadpoles, a scale for measuring in milligrams, and a ruler for length measurements in centimetres. As previously described, before observations, the camera setup was prepared to ensure optimal recording conditions and minimize interaction with the subjects during the acclimation and assay periods. Tadpoles were randomly captured using a small mesh dip net, with five individuals selected from the target mesocosm and placed in a temporary container filled with mesocosm water. These containers were then carefully transported to the behavioural assay tent at the study site. Tadpoles were transferred into their respective behavioural arenas using the dip net. The behavioural assay was conducted as previously described.

After recording, tadpoles were returned to the temporary containers pre-filled with mesocosm water. Subsequently, tadpoles were measured, weighted, and photographed from the dorsal view for all samples. Depending on the subsampling schedule, tadpoles were either euthanized, apical endpoints sampled, and then preserved on dry ice, or were released back into

their original mesocosm. These behavioural assays were designed to be minimally invasive, ensuring minimal stress to the tadpoles upon release back into the corresponding mesocosms.

In addition to the above-described assessments of basal behaviour, during the final day of behavioural assessments (study day 28), an additional behavioural test on a new subset of tadpoles from the tanks was conducted where a predatory chemosensory cue was added to further determine predatory avoidance behaviours in the tadpoles (Reeves *et al.*, 2011; Mirza *et al.*, 2013). The predator cue consisted of water that had been exposed to the existing predacious diving beetle; a known predator at the study site, in the family Dytiscidae. Tadpoles were observed being predated on by these diving beetles in their natural habitat (the ponds where the egg masses were collected from) so it was appropriate to include this as the predatory cue for the final behavioural assay (Reeves *et al.*, 2011). All processing and analysis of the recordings was conducted as previously described in section 3.2.6 *Behavioural Response Assessments*.

3.2.7 Statistical Analysis

All data analysis was conducted in R version 4.2.2 (R Core Team 2024). Project set up and basic data wrangling was conducted using the following packages: here v. 1.0.1(Müller, 2020), janitor v. 2.2.0 (Firke, 2023), knitr v. 1.43 (Xie, 2014, 2015, 2023), lme4 v. 1.1.34 (Bates *et al.*, 2023), multcomp v. 1.4.25 (Genz *et al.*, 2023), multcompView v. 0.1.10 (Graves *et al.*, 2024), patchwork v. 1.1.2 (Pedersen, 2022), rmarkdown v. 2.27 (Xie *et al.*, 2018; Xie *et al.*, 2020; Allaire *et al.*, 2023) and tidyverse v. 2.0.0 (Wickham *et al.*, 2019). Model outputs were formatted using broom.mixed v. 0.2.9.5 (Robinson *et al.*, 2023). All figures were created using ggplot2 from tidyverse v. 2.0.0 (Wickham *et al.*, 2019) with additional functions from patchwork v. 1.2.0 (Pedersen, 2022), scales v. 1.3.0 (Pedersen, 2022; Wickham and Seidel, 2022), and viridis v. 0.6.5 (Garnier, 2023).

3.2.7.1 General Water Quality

To evaluate the effects of treatments on the analysis of general water quality parameters, a one-way ANOVA was performed for each general water quality parameter. The experimental design involved treatment groups as a categorical independent variable, with parameter

concentrations being the dependent variables. A linear model (lm()) was fitted using mesocosm treatment groups as the predictor variable. The model tested whether mean parameter concentrations differed significantly among the mesocosm treatments. In cases where the ANOVA indicated significant differences, post-hoc multiple comparisons were performed using least-squares means (lsmeans). This allowed for pairwise comparisons between treatment groups to determine which specific groups differed significantly from each other. All comparisons were tested at a 95% confidence level, and significance was determined using p-values adjusted for multiple comparisons.

3.2.7.2 Exposure Characterization

To assess the average exposure of the tadpoles to oil-derived hydrocarbons throughout the exposure period, water tPAC data from the various analytical sampling events was used to calculate Time Weighted Average (TWA) concentrations for the treatments. Due to the complexity of the study, it was not possible to pair sampling for tPAC with tadpole sampling. As such, the two sampling events for tadpole endpoints fell in between tPAC sampling days. To estimate the TWA concentration at the time of the two tadpole sampling events, a LOESS model was fit to the TWA concentration data for the tPAC analytical sampling days using the loess function from the basic stats package, leaving all default settings. TWA concentrations for the tadpole sampling events in the treatments were then predicted from these models with the predict function from the stats package. To compare the TWA, the CHV-WAF water data were grouped by exposure day, and for each group, a linear model was fitted to compare TWA concentrations across treatment groups. Specifically, TWA concentrations at each sampling time were modelled as a function of the treatment group. Model fits were summarized, and statistical significance was determined at a p-value of 0.05. Multiple comparisons between treatments were conducted using least-squares means (lsmeans), with pairwise contrasts computed to identify differences between groups. This method of multiple comparisons is robust and is similar to a traditional Tukey's post-hoc test.

3.2.7.3 Tadpole Endpoints

3.2.7.4 Growth Rate

A logistic growth model was used to estimate the growth rate of the tadpoles (based on weight) and associated parameters over the study period. Measurements collected from tadpoles used for behavioural analysis were used as these were the only tadpoles which were measured regularly throughout the study. To make use of tadpole-specific data within the model a mixed effects model approach was used where individual tadpole data were fed into the model and tank (replicate) was used as a random effect. For this, the nlme function from the nlme package (José Pinheiro *et al.*, 2024) was used to fit a logistic growth model via the SSlogis self-starting function. The random effect of 1|tank was applied only to the Asymptote. The different model parameters such as k (maximum weight (in grams)) and b (intercept) and r (growth rate per day (in grams)), were estimated with their associated 95% level of confidence.

3.2.7.5 Apical Endpoints

Tadpole apical endpoints were measured individually and for each of the sampling events separately. To be able to use individual tadpole data (i.e. avoid condensing to one value per tank) without incurring pseudo-replication, tadpole biological endpoints were analyzed via a mixed effects model approach with the presence or not of oil as the main treatment variable and tank as a random effect. The lmer function within the lme4 package Bates *et al.* (2015) was employed to fit these models using the restricted maximum likelihood (REML) option. Conformity to model assumptions was assessed by visual inspection of the distribution of residuals. Pairwise multiple comparisons of the treatments were conducted using the lsmeans function from the emmeans package Lenth (2024), followed by the cld function from the multcomp package (Hothorn *et al.*, 2008).

For each combination of Gosner stage and morphological endpoint, a linear mixed-effects model (LMM) was fitted using the treatment group as the fixed effect and the mesocosm as a random effect to account for the clustering of tadpoles within each tank. The models were fitted using Restricted Maximum Likelihood (REML) to estimate the parameters. After fitting the

models, the results were summarized to extract coefficient estimates, standard errors, t-values, and p-values, providing a detailed overview of the effect of treatments on each morphological parameter. To ensure the validity and normality of the models, diagnostic checks were performed. A Q-Q plot was generated to assess the normality of residuals, while a histogram of residuals and a plot of residuals versus fitted values were used to evaluate the distribution and homoscedasticity of the residuals.

3.2.7.6 Bioaccumulation of Total Polycyclic Aromatic Compounds

Tissue tPAC concentration data was simplified to one single value per tank (n = 3 per treatment) by averaging the tPAC tissue concentration data for all tissue pools collected from the same treatment. Differences in tissue tPAC concentrations between treatments were assessed via analysis of variance (ANOVA) using the lm() function within R, followed by the anova() function. Assumptions of the ANOVA test were visually inspected by exploring the distribution of residuals. Pairwise multiple comparisons of the treatments was conducted using the lsmeans function from the emmeans package Lenth (2024), followed by the cld function from the multcomp package (Hothorn et al., 2008).

3.2.7.7 Behavioural Assays

For the reference-toxicant tests with caffeine and ethanol, individual tadpole values were averaged to a single mean replicate value. This was needed to simplify the fitting of non-linear concentration-response models. Concentration-response models were fit using the drc package in R (V3.0) (Ritz, 2010; Ritz *et al.*, 2016). The fit of two different models was attempted for each compound/endpoint combination: one was a standard 4-parameter log-logistic model (fitted as the LL.4() function), and the other was a Cedergreen model able to account for hormesis, which was apparent in some of the responses (Cedergreen *et al.*, 2005; Belz *et al.*, 2012). Additionally, the Cedergreen model was run using six different values of alpha (0.10, 0.25, 0.36, 0.50, 0.75, and 1.00), the parameter which defines the reflecting the steepness of the hormesis peak. This yielded a total of 7 different models that were tested for each compound/endpoint combination. Upper and lower bounds of the concentration-response curve were set for each individual endpoint based on the mechanistic understanding of the expected response as well as preliminary observations of the

data. This way, lower bounds of 0 were set for the total distance moved and the mean nearest neighbour distance. An upper bound of 100 (% of the time) was set for the time spent in the inner half of the arena.

The best fitting model among the seven tested models was selected using a combination of visual assessment of the fit and the value of the Akaike Information Criterion (AIC) value. The concentration causing a 50% effect (EC_{50}) was calculated from the selected model using the ED() function from the drc package. In general, no concentration-response models were fit to cases where chemical exposure did not result in at least a 50% effect (increase or decrease).

For the main study, tadpole behavioural endpoints were assessed individually and for each of the sampling events separately. To be able to use individual tadpole data (i.e. avoid condensing to one value per tank) without incurring pseudo-replication, tadpole behavioural endpoints were analyzed via a mixed effects model approach with the presence or not of oil as the main treatment variable and tank (end by extension behavioural arena) as a random effect. The lmer function within the lme4 package; Bates et al. (2015), was employed to fit these models using the restricted maximum likelihood (REML) option. Conformity to model assumptions was assessed by visual inspection of the distribution of residuals. Pairwise multiple comparisons of the treatments were conducted using the lsmeans function from the emmeans package, Lenth (2024) followed by the cld function from the multcomp package (Hothorn et al., 2008). To ensure the validity and normality, of the models, diagnostic checks were performed. A Q-Q plot was generated to assess the normality of residuals, while a histogram of residuals and a plot of residuals versus fitted values were used to evaluate the distribution and homoscedasticity of the residuals.

3.3 Results

3.3.1 General Water Quality

High-frequency temperature data recorded in the mesocosms from May 30, 2021, to July 30, 2021 showed no significant temperature (°C) differences among treatments ($p = 0.24$) (SI Table 3-8, SI Table 3-9, SI Table 3-10), however temperatures showed only slight differences among the treatments. The no oil treatment had an average temperature of 21.1°C (± SD 4.1), while the oil

with plants and oil without plants treatments had mean temperatures of 21.5°C (± SD 4.0) and 21.9°C (± SD 4.1), respectively (SI Table 3-9). Multiple comparisons could not be evaluated due to missing logger data across some treatment replicates (SI Table 3-10).

Dissolved oxygen (DO), pH, specific conductivity across treatments from May 30, 2021, to July 30, 2021 revealed that the only significant differences when comparing parameters between treatments was in dissolved oxygen percentages (between the no oil and oil without plants treatments) while other parameters like pH, specific conductivity did not exhibit statistically significant differences among treatments (SI Table 3-11, SI Table 3-12, SI Table 3-13). The oil without plants treatment had the highest DO concentration in mg/L (11.3 ± SD 14.1 mg/L), compared to the no oil treatment (8.0 ± SD 1.0 mg/L) and the oil with plants treatment (8.3 ± SD 0.8 mg/L) (SI Table 3-13).

Water nutrient parameters such as alkalinity, ammonia, calcium, chlorine, dissolved organic carbon (DOC), and total dissolved nitrogen (TDN) were measured between May 30, 2021, and July 30, 2021 (SI Table 3-14). Ammonia was the only parameter to display statistically significant differences between the treatments, specifically between the oil with plants and oil without plants treatments (multiple comparisons $P = 0.04$). Ammonia concentrations were greater in the oil without plants treatment (90.21 µg/L ± SD 93.12) compared to no oil (74.43 µg/L ± SD 103.10) and oil with plants (45.29 µg/L ± SD 71.95) (SI Table 3-15). Other nutrient parameters did not exhibit significant differences between treatments (SI Table 3-16).

3.3.2 Water Accommodate Fraction Chemistry

The temporal distribution of tPAC concentrations is presented in Figure 3-1. A clear increase on tPAC can be seen starting on day 0 (WAF application day). Time weighted average (TWA) concentrations of total polycyclic aromatic compounds (tPACs) were calculated for the average exposure day associated to each of the main sampling events (Figure 3-2, SI Table 3-17).

On Day 32, TWA concentrations for mesocosms with no oil ranged from 38.7 to 69.5 ng/L, while mesocosms with oil and plants showed concentrations from 49.0 to 66.1 ng/L. Oil without plants mesocosms had a wider range, with concentrations between 57.3 and 111.4 ng/L. On Day 36, TWA concentrations followed a similar pattern, with no oil treatments ranging from 36.7 to

67.6 ng/L, oil with plants from 46.3 to 62.5 ng/L, and oil without plants from 54.3 to 111.6 ng/L (Table 3-1). The multiple comparison test on TWA concentrations revealed no statistically significant differences between the treatments (oil without plants, oil with plants, and no oil) on either exposure day (Figure 3-2, SI Table 3-17).

3.3.3 Tadpole Endpoints

3.3.3.1 Survivorship

The greatest survival of tadpoles occurred in the oil without plants treatment, with a mean survival rate of 92.0% (± SD 2.6%). Conversely, the oil with plants treatments exhibited the lowest survivorship, with a mean survival rate of 64.2% (± SD 15.6%). The no oil condition had survivorship with a mean survival rate of 89.9% (± SD 9.9%). These survivorship values were important to show if tank conditions were sufficient for tadpole survival and section 3.4 will extrapolate why this survivorship was experienced. Volume of the tanks and therefore tadpole density per litre were relatively consistent across all treatment conditions, indicating that these factors did not significantly influence the survival outcomes (Table 3-2).

3.3.3.2 Apical Endpoints

The F-statistics indicate differences in variability among the different apical endpoints measured in each GS class (SI Table 3-18). The multiple comparisons test for both sampling periods (GS classes 34-36 and 40-42) on endpoints of weight, total length, snout-vent length and Gosner stage did not display statistically significant differences across the treatments; oil without plants, oil with plants, and no oil with plants (Table 3-3, SI Table 3-19).

Gosner stage did not show significant differences between treatments and sampling events (Table 3-3, Figure 3-3). For the first sampling event, p-values were 0.85 and 0.60, when comparing the control treatment to the oil with plants and oil without plants treatments, respectively, and a p-value of 0.91 between the oil with plants and oil without plants treatments. For the second sampling period (GS stage 40-42), Gosner stage comparisons yield p-values of 0.69, 1.00, and 0.67 (SI Table 3-18, SI Table 3-19).

Snout vent length did not show significant differences between treatments and sampling events (Table 3-3, Figure 3-3). Snout-vent length comparisons also show no significant differences, with p-values of 0.52 and 0.87, when comparing the control treatment to the oil with plants and oil without plants treatments, respectively, and a p-value of 0.77 between the oil with plants and oil without plants treatments. For the second sampling period (GS class 40-42), snout-vent length comparisons result in p-values of 0.77, 0.46, and 0.21 (SI Table 3-18, SI Table 3-19).

Total length did not show significant differences between treatments and sampling events (Table 3-3, Figure 3-3). Total length comparisons also show no significant differences, with p-values of 0.69 and 0.87, when comparing the control group to the oil with plants and oil without plants treatments, respectively, and a p-value of 0.91 between the oil with plants and oil without plants treatments. For the second sampling period (GS class 40-42), total length comparisons show p-values of 0.99, 0.48, and 0.56 for the same treatment comparisons (SI Table 3-18, SI Table 3-19).

Weight did not show significant differences between treatments and sampling events (Table 3-3, Figure 3-3). For the first sampling period, GS class 34-36, weight comparisons show no significant differences with p-values of 0.49 and 0.93 when comparing the control treatment to the oil with plants and oil without plants treatments, respectively, and a p-value of 0.66 between the oil with plants and oil without plants treatments. Similarly, for the second sampling period (GS class 40-42), weight comparisons show no significant differences with p-values of 0.73, 0.40, and 0.16 for the same treatments comparisons listed above (SI Table 3-18, SI Table 3-19).

3.3.3.3 Growth Rate

Growth rate was determined from weekly measurements of tadpoles (Table 3-4). est_k represents the growth rate across treatments, est_b represents the growth rate changes over time and est_r represents the overall growth of the tadpoles. The no oil (control) mesocosms had medium and stable growth rate in comparison to the other calculated rates (est_k = 0.805, 95% CI: 0.727 to 0.882). The coefficient b was (est_b = 3.297, 95% CI: -0.418 to 7.011) and the coefficient r was (est_r = 0.174, 95% CI: 0.061 to 0.288).

The oil with plants mesocosms showed the greatest growth rate (est_k = 0.929, 95% CI: 0.676 to 1.182). The coefficient b was (est_b = 2.272, 95% CI: 0.113 to 4.430) while the coefficient r was (est_r = 0.119, 95% CI: -0.011 to 0.250).

The oil without plants mesocosms displayed the lowest growth rate (est_k = 0.691, 95% CI: 0.566 to 0.816), however, oil without plants treatment had the highest growth rate changes over time (est_b = 3.725, 95% CI: 0.617 to 6.833) compared to the oil with plants treatment, while the coefficient r (est_r = 0.165, 95% CI: 0.066 to 0.265) was comparable to the control treatment (Figure 3-4, Figure 3-5, Table 3-4).

3.3.3.4 Histology

There was no apparent difference among treatment groups in the histological appearance of the liver (Figure 3-6, SI Table 3-2). The liver parenchyma was comprised of cords of relatively small hepatocytes interspersed with sinusoids and larger blood vessels containing erythrocytes. The hepatocytes contained spherical dark-staining nuclei and small round vacuoles in the cytoplasm.

3.3.4 Bioaccumulation of Total Polycyclic Aromatic Compounds

The analysis of total polycyclic aromatic compounds in *Lithobates sylvaticus* tadpole tissues reveals differing concentrations based on GS class and treatments. The mean tPAC concentration within tadpole tissues within the first sampling event (GS 34-36) ranged from 0.41 ng/g lipid (± SD 0.70) in the no oil treatment to 0.82 ng/g lipid (± SD 0.72) in the oil without plants treatment. The oil with plants treatment had a mean tPAC concentration of 0.51 ng/g lipid (± SD 0.85). As for the second sampling event (GS 40-42), mean concentrations of tPAC in tadpole tissue were lower, with the no oil treatment showing 0.16 ng/g lipid (± SD 0.23), oil with plants at 0.33 ng/g lipid (± SD 0.29), and oil without plants at 0.02 ng/g lipid (± SD 0.02) (Table 3-5).

No significant differences were observed for any of the Gosner stage classes analyzed. For the first sampling event (GS 34-36), the treatment effect on tPACs was not statistically significant, with an F-statistic of 0.241 and a p-value of 0.793. For the second sampling event (GS 40-42), the treatment effect was also non-significant, with an F-statistic of 1.575 and a p-value of 0.282,

indicating that the different treatments (no oil, oil with plants, and oil without plants) did not lead to significant differences in tPAC concentrations in tadpole tissues.

Conducting a multiple comparisons test, the statistical analysis of tPAC bioaccumulation within tadpole tissues of the experiment and Gosner stages revealed no significant differences between the treatments. In the first sampling period, the comparison between no oil and oil with plants had an estimate of -0.099 with a p-value of 0.986. When looking at the comparison between no oil and oil without plants, the estimate was -0.411 with a p-value of 0.791. When looking at the comparison between oil with plants and oil without plants, the estimate was -0.312 with a p-value of 0.872 (Figure 3-7, SI Table 3-20, SI Table 3-21, SI Table 3-22). For the second sampling timepoint for tPAC bioaccumulation within tadpole tissues of the experiment across different treatments and Gosner stages, the comparison between no oil and oil with plants had an estimate of -0.168 with a p-value of 0. 624.When looking at the comparison between no oil and oil without plants, the estimate was 0.141 with a p-value of 0.771. When looking at the comparison between oil with plants and oil without plants, the estimate was 0.308 with a p-value of 0.256 (Figure 3-7, SI Table 3-20, SI Table 3-21, SI Table 3-22).

3.3.5 Behavioural Assays

3.3.5.1 Reference Toxicant Behavioural Assay

Exposure to caffeine up to 10 mg/L did not cause noticeable changes in socialization (mean distance to nearest-neighbour) or thigmotaxis (time spent in the inner area of the arena) (Figure 3-8). As such it was not possible to fit a concentration-response model and derive EC_{50} values, exposure to caffeine did elicit concentration-dependent changes in total distance travelled. A hormetic response was apparent, which was confirmed by the best fit of a Cedergreen model (p-value = 0.25). The lower tested concentrations resulted in an increase in total distance moved up to ~700 cm (compared to an average ~300 cm for the control treatment), increasing concentrations resulted in an eventual decrease in total distance moved with an EC_{50} extrapolated to be 46.20 mg/L (higher than the highest tested concentration of 10 mg/L) (Figure 3-8).

Exposure to ethanol up to 2.68% (v:v) resulted in higher than 50% effects on both total distance moved and time spent in the inner area of the arena. Ethanol exposure up to 2.68 % showed a concentration-response pattern in mean nearest neighbour distance with a reduction up to ~40% from control values (Figure 3-8). Given the apparent trend in the mean nearest-neighbour distance, a concentration-response model to all three endpoints was fit. A simple log-logistic model was the best fit for time spent in the inner half of the arena (as %) and mean nearest-neighbour distance although with the responses showing different directions (decreasing mean nearest-neighbour distance and increasing time spent in the inner area of the arena with increasing ethanol concentration). Estimated EC_{50}s were 30.86 % (extrapolated from the maximum value tested of 2.68 %) for mean nearest-neighbour distance and 2.68% for time spent in the inner area of the arena. Similarly to what was observed for caffeine, low ethanol exposures resulted in increased total distance moved (up to ~800 cm vs ~300 cm for the control), with distance declining sharply in higher exposures. As such, a Cedergreen model (alpha = 0.25) accounting for hormesis resulted in the best fit. The EC_{50} for the total distance moved was estimated to be 1.40 % ethanol.

3.3.5.2 Oil-Exposed Behavioural Assays

There were no significant effects of treatment on total distance or nearest neighbour distance at any study day (Figure 3-9, Figure 3-10). For Day 7, the total distance showed an F-statistic of 1.292, nearest neighbour distance had an F-statistic of 0.606 and percent time spent in inner sector (%) has an F-statistic of 1.835. On Day 15, total distance had an F-statistic of 0.14 (p-value = 0.868), nearest neighbour distance had an F-statistic of 0.058 and percent time spent in inner sector (%) has an F-statistic of 0.188. On Day 21, total distance was had an F-statistic of 0.663, nearest neighbour distance had an F-statistic of 2.031and percent time spent in inner sector (%) has an F-statistic of 2.159. On Day 28, total distance had an F-statistic of 1.35, nearest neighbour distance an F-statistic of 1.771, and percent time spent in inner sector (%) has an F-statistic of 0.809 (SI Table 3-23).

Multiple comparisons model did not reveal significant differences between treatments for either total distance, nearest neighbour distance, or percent time spent in inner sector at any study day. On Day 7, comparisons between treatments yielded p-values ranging from 0.426 to 0.995 for total distance and p-values from 0.573 to 0.965 for nearest neighbour distance and p-values ranging

from 0.33 to 0.983 for percent time spent in inner sector. For Day 15, p-values ranged from 0.863 to 0.977 for total distance and from 0.939 to 0.986 for nearest neighbour distance and p-values ranged from 0.819 to 0.955 for percent time spent in inner sector. On Days 21 and 28, all comparisons showed p-values above 0.2 for both nearest neighbour distance, total distance, and percent time spent in inner sector, indicating no significant treatment effects (SI Table 3-24).

When examining the results when predator cue was present on Day 28, total distance (with predator cue) had an F-statistic of 0.456, nearest neighbour distance (with predator cue) has an F-statistic of 0.488, and percent time spent in inner sector (%) (with predator cue) had an F-statistic of 1.633 (Table 3-6, SI Table 3-23). Multiple comparisons showed no significant differences between treatments for percent time spent in inner sector, total distance, or nearest neighbour distance. The p-values for these comparisons ranged from 0.548 to 0.999, indicating no significant effects of predator cue on the behaviour metrics between the treatments (Figure 3-9, SI Table 3-24).

When examining the multiple comparison results of addition of a predator cue on Day 28 to results from multiple comparisons of behavioural assays also conducted on Day 28 (without a predator cue addition), it revealed that total distance in the no oil and oil with plants mesocosms were statistically significant. Predator cue addition compared to the basal behaviour, caused a higher total distance travelled. Additionally, a second behavioural assay endpoint; nearest neighbour distance in the oil with plants and oil without plants was statistically significant when comparing against basal behaviour. This time, however, tadpoles decreased nearest neighbour distance with the addition of a predator cue (Figure 3-10, Table 3-6).

3.4 Discussion

Conducted in 2021 as part of the Floating Wetland Treatments to Enhance Remediation (FLoWTER) study at the International Institute for Sustainable Development-Experimental Lakes Area in Ontario, Canada, this study investigated the impact of chronic exposure to CHV-derived hydrocarbons (WAF) on wood frog tadpoles. The primary objective of the study was to characterize how chronic exposure to CHV-derived hydrocarbons (WAF) affects the growth, development, histopathological changes, and bioaccumulation in wood frog tadpoles. Additional

objectives included investigation on how exposure to CHV-derived hydrocarbons (WAF) influenced tadpole behaviour and, lastly how to examine whether the presence or absence of EFWs impacts the growth, development, and behaviour of wood frog tadpoles exposed to CHV-derived hydrocarbons (WAF). Overall, it was found that even at the highest total polycyclic aromatic hydrocarbon (tPAC) concentrations present in the study, no significant impact was observed on the endpoints measured.

When comparing water quality parameters across treatments and control mesocosms, most showed no significant differences. However, ammonia (NH_3) levels were notably higher in the oil without plants mesocosms compared to both the no oil and oil with plants treatments, with significant differences observed between the oil with plants and oil without plants. Jofre *et al.* (2000) examined the effects of unionized NH_3 concentrations up to 2000 µg/L on green frogs and leopard frogs, finding that green frogs exposed to as low as 600 µg/L experienced reduced survival and growth. These concentrations, however, are approximately six times higher than the maximum ammonia levels in this study's mesocosms (maximum concentration of 90.21 µg/L ± SD 93.12), suggesting that adverse effects on tadpoles are unlikely. Measurements for dissolved oxygen (DO) percentages in the oil without plants mesocosms had the highest DO levels compared to the other treatment (oil with plants) and control (no oil). This suggests that oil presence, particularly without vegetation, can lead to elevated DO levels, possibly due to decreased biological oxygen demand from plants and microbial activity. It is possible the planted engineered floating wetlands contributed oxygen through photosynthesis but also had microbial activity that was consuming oxygen, which may explain the reduced DO in the mesocosms with planted EFWs (Afzal *et al.*, 2019; Rehman *et al.*, 2019).

A similar study conducted at IISD-ELA in 2019 that exposed tadpoles to differing concentrations of diluted bitumen in *in-situ* freshwater enclosures, measuring similar developmental parameters found that wood frog tadpole total lengths ranged from 43.2- 51.3 mm and 0.88–1.25g for weight (Patterson *et al.*, 2022). In comparison to this 2021 study, the total length of the tadpoles across the three treatments during the second sampling event was smaller and not in the range experienced in the study (10.7% to 16.2% shorter compared to the 2019 study) and 15.5% to 34.3% lighter compared to the 2019 study, with the largest impact in the oil without plants treatment (Patterson *et al.*, 2022)(Table 3-3). This could indicate that aquatic plants' root

structure in oil with plants treatments provided benefits for tadpoles, allowing tadpoles to allocate more energy toward growth and grazing despite potential environmental stressors from the CHV-derived hydrocarbons (WAF). As it is known that when tadpoles are exposed to higher PACs, they must allocate a portion of their limited energy reserves toward producing detoxification enzymes (Hersikorn and Smits, 2011). This energy diversion reduces the resources available for other essential physiological functions, potentially affecting their fitness and survival (Hersikorn and Smits, 2011).

No significant differences were observed in apical endpoints (weight, total length, snout-vent length, Gosner stage), the findings in this study align with previous research on wood frogs, which suggests that the water-soluble components of oil are relatively less harmful to amphibians (Hedtke and Puglisi, 1982). Past research indicates that polycyclic aromatic compounds (PACs) are often efficiently metabolized in wood frog tadpoles, though their toxicological effects are not always fully understood or characterized (Bilodeau *et al.*, 2019; Mundy *et al.*, 2019).

Tadpoles in the mesocosms with no oil exhibited the highest and most stable growth rate constant in the comparison between the tadpoles of other treatments, indicating optimal growth conditions typical of a clean environment, with the exception of ammonia (OECD, 2009, 2015). Now, examining the oil with plants tanks, the results showed the highest growth rate constant among the oil-exposed tanks, suggesting that the presence of plants may help mitigate some of the negative effects of oil exposure, thus providing favourable growth conditions. However, while there was a positive effect on growth with the presence of planted EFWs, it was less pronounced than in the no oil mesocosms, and tadpole growth remained less consistent compared to the control treatments. The oil without plants group experienced the lowest overall growth rate constant, further revealing the combined impact of CHV WAF and the impacts of the absence of plants on tadpole growth. Results from a study with tadpoles reared in fresh oil sands process-affected materials (OSPM) area showed that tadpoles take significantly longer to complete metamorphosis (or may not complete it at all), compared to control or older OSPM-affected areas (Hersikorn and Smits, 2011; Thambirajah *et al.*, 2019).

The lowest growth rate experienced in the oil without plants treatment may also reflect the lack of structural refuges and altered resource availability in the unplanted treatments, suggesting

that plants can buffer contaminants' impact on amphibian growth. This result shows the presence of plants potentially helped tadpole growth and the potential impact of limitation of food availability as well. Although tadpoles within the oiled unplanted tanks appeared smaller in many apical endpoints, these were not statistically significant, indicating that there was enough food available in the systems not to experience detectable reduced development on any endpoints. If significant differences in tadpole apical endpoints and metamorphic stages in addition to growth rate, as measured by Gosner stages, had been observed, this could have indicated limited food availability. According to literature, food deprivation in wood frogs and similar species leads to developmental delays, although weight and SVL remain statistically similar to those not food deprived (Bulaeva et al., 2015). Bulaeva *et al.* (2015) restricted food for wood frogs before pre-metamorphosis and found that reduced feeding/food availability causes delays in tadpole development which, then lasts to metamorphosis. The present study can be comparable for the fact that wood frog tadpoles began the exposure period within the mesocosms at GS 22.9 +/- SD 1.6 (on May 31st, 2021), early enough in their pre-metamorphic development to experience delayed metamorphosis if experiencing food restricting conditions (Bulaeva *et al.*, 2015). An earlier study by Audo *et al.* (1995) also found that a 10-day period of food restriction during early stages (~Gosner stage 25) in *Hyla chrysoscelis* tadpoles delayed metamorphosis, however the restricted tadpoles' size and weight were comparable to those of the control group (Audo *et al.*, 1995). This suggests that tadpoles experiencing early food restrictions may eventually catch up in body size (length and weight), but they still face challenges later on due to delayed metamorphosis (Bulaeva et al., 2015).

Perhaps, the findings from this study are more of an indication of the plants impact on PACs, and not on food availability. CHV WAF (although minimal) exposure negatively impacted tadpole growth rates, but the presence of plants partially alleviated these effects, (although not significantly), highlighting the potential role of engineered floating wetlands in mitigating CHV-derived hydrocarbons (WAF) toxicity in impacted freshwater aquatic environments in spill event scenarios.

Tadpoles exposed to CHV WAFs experienced no cellular damage to the liver as assessed via histopathological analysis. The absence of histological alterations or cellular deformities in this study could reflect specific mesocosm conditions that limited CHV's bioavailability or toxicity.

Thambirajah *et al.* (2019) highlighted that amphibian developmental impacts from hydrocarbon contaminants are often moderated by environmental factors and may differ based on the compound's characteristics, concentration, and exposure duration. The low overall bioaccumulation levels further suggest limited PAC uptake in tadpole tissues under the experimental conditions and that the concentrations of tPACs in tadpole tissues across treatments showed no significant differences between treatments in either sampling period. However, while no significant treatment effect was observed, the slightly higher tPAC levels in the oil without plants treatment during the initial stages indicate a potential for greater PAC exposure in plant-absent tanks, possibly due to reduced adsorption or degradation by plants. In all, regardless of developmental stage and sampling event, the tPAC concentrations in tadpole tissues were unaffected by the presence or absence of CHV-WAF and planted engineered floating wetlands.

While the impact of engineered floating wetlands on growth and apical endpoints remains speculative, survivorship data showed that tadpoles in the oil without plants treatment had the highest survival rate, with the oil with plants treatment exhibiting lower survival. In this study, tadpole survival rates varied with oil exposure conditions, with the highest survival rate of 92.0% recorded in the oil-without-plants mesocosm, where tPAC levels ranged from 57.3 to 111.6 ng/L on day 32 and from 54.3 to 111.6 ng/L on day 36. The oil with plants treatment showed a reduced survival rate of 64.2%, with tPAC concentrations from 49.0 to 66.1 ng/L on day 32 and from 46.3 to 62.5 ng/L on day 36. The control (no oil) treatment had a high survival rate of 89.9% at tPAC levels of 38.7 to 69.5 ng/L on day 32 and 36.7 to 67.6 ng/L on day 36. Lara-Jacobo *et al.* (2019) comparatively used chemically enhanced WAFs with much higher tPAC levels (152,000-247,000 ng/L tPACs) and that resulted in significantly higher mortalities (100% mortality). When tadpoles in that experiment were exposed to lower concentrations (700, 3,800–8,200 ng/L), the concentrations resulted in survival rates of 96%, 92% and 91.5%, respectively, demonstrating acute toxicity at these levels (Lara-Jacobo et al., 2019). This current tank study, using substantially lower tPAC levels (up to 111.6 ng/L), recorded survival rates from 64.2% (oil with plants), 89.9% (no oil) to 92% (oil without plants mesocosm), suggesting similar survivorship to Lara-Jacobo *et al.* (2019) except in the oil with plants mesocosms. The reason behind reduced survival in the oil with plants mesocosms may only be due to the detection rate calculated as the unknown fate, which accounts for tadpoles that were unaccounted for during the sampling period and, if the mortality event was not definitive, was highest in the oil with

plants condition. This suggests that a significant proportion of tadpoles in this treatment were neither found alive nor confirmed dead, indicating the impact of environmental conditions on detectability. For example, with the presence of plant root biomass and detritus leaf material, tadpoles were effective in hiding and avoiding detection during sampling selection events, and detection of tadpoles was a lot easier in the tanks without planted EFWs as the quantity of biomass was visually a lot less than other treatments. Another contributing factor to reduced survival in oil with plants mesocosms was that tadpole predators became apparent within the tanks throughout the field experiment, thus potentially having a mortality event without being detected and accounted for, as Dytiscidae and *Anisoptera infraorder* (dragonfly larvae) are efficient tadpole predators. Additionally, the presence of vegetation might create more complex microhabitats, leading to higher exposure to sediments potentially containing PACs or decreased avoidance behaviour in tadpoles.

When it comes to the assessed behavioural endpoints, the reference-toxicant test confirmed that our methodologies were able to detect changes in tadpole behaviour after acute chemical exposure. Both caffeine and ethanol displayed a hormetic concentration response in total distance moved, with caffeine not displaying changes in socialization and/or thigmotaxis (Carlson and Langkilde, 2013). Ethanol, however, did induce changes in socialization (reduced distance to the nearest neighbour) and thigmotaxis (increased time spent in open areas), which appear consistent with its expected anxiolytic (Carlson and Langkilde, 2013).

Despite the reference-toxicant test's validation of our behavioural assessment methodologies, chronic exposure to oil-derived hydrocarbons in the main study resulted in minimal differences in tadpole activity, distance travelled, and time spent in thigmotaxis across treatments when tested in the absence of predator cues. The introduction of a predator cue, as tested on day 28, resulted in behavioural changes (when compared to behaviour in the absence of the cue), which were endpoint and treatment-specific. Specifically, the introduction of the predator cue increased the total distance travelled in the no oil reference and the oil with plants treatment. The usual antipredator behaviour of larval wood frog tadpoles involves a reduction in activity upon detection of predation cues (Ferrari *et al.*, 2008; Reeves *et al.*, 2011). However, in our study, predator cue addition resulted in statistically significant increases in total distance travelled within these mesocosms. Furthermore, another behavioural assay endpoint—nearest neighbour distance—also showed significant treatment-specific responses. In the oil with plants

and oil without plants treatments, tadpoles decreased their nearest neighbour distance with the addition of a predator cue (Figure 3-10, Table 3-6). This increase is not consistent with previous research, but increased activity was not observed in the oil without plants treatment (Ferrari et al., 2008; Reeves et al., 2011; Ågerstrand et al., 2020). The introduction of the predator cue also resulted in a significant increase in socialization (reduction in mean nearest neighbour distance), which was not observed in the not oiled reference treatment. At this time, we do not have a mechanistic explanation for this altered behaviour in exposed tadpoles, but it could be the case that the presence of oil may influence tadpole anti-predator responses, potentially due to altered perception or stress-related behaviour in contaminated environments (Reeves *et al.*, 2011; Mirza *et al.*, 2013; Urszán *et al.*, 2015; Ågerstrand *et al.*, 2020).

3.5 Conclusions

Wood frog tadpoles, recognized as sensitive indicators of environmental contaminants and toxicity, did not exhibit significant disruptions in apical developmental markers such as Gosner stage, snout-vent length, total body length, or weight during the two sampling events. Despite hypothesizing potential developmental impacts and changes in tadpole activity due to CHV-WAF exposure, the study generally found no significant treatment effects on growth parameters, bioaccumulation of oil-derived compounds, or behavioural responses, including in the presence of a predator cue related to characterizing effects of PACs toxicity on wood frog tadpoles. This lack of discernible impact across the examined endpoints suggests that the minimal levels of tPACs (total polycyclic aromatic compounds) present in the mesocosms after the addition of the CHV WAF did not critically affect the early developmental stages of wood frogs. This outcome may be attributed to the low levels of tPACs present in the exposure environment and the efficient metabolization of certain PACs by tadpoles, which may have minimized toxicological effects. Exposure levels after a few days were minimal, and therefore, when understanding how tadpoles have been documented to efficiently metabolize certain PACs effectively, the concentrations the early life stage tadpoles were exposed to were minimal to warrant any observed toxicological effects.

Behavioural assays, however, did detect expected responses in the ethanol (anxiolytic) and caffeine (anxiogenic) reference-toxicant behavioural tests, confirming the ability of our assay to

quantity tadpole behaviour. Therefore, if tPAC concentrations had been higher throughout the exposure period, this methodology would have likely been able to detect responses, reinforcing the potential utility of behavioural assays in future toxicity studies.

The findings from this study do, however, provide insights into the characterization of chronic effects of conventional heavy crude oil (CHV) exposure on wood frog (*Lithobates sylvaticus*) tadpoles. These findings emphasize the ecological significance of plant presence in oil-contaminated freshwater systems. While plants and root systems may offer structural habitat, they also appear to modify tadpole behaviour and survival, potentially altering the ecological consequences of oil exposure. The ability to observe physiological, behavioural, histological and bioaccumulation responses reveal the potential for CHV contamination to disrupt freshwater ecosystems at certain concentrations. Given the crucial role of amphibians in ecosystem functioning and their status as environmental bioindicators, these results underscore the need for further research to fully understand the long-term impacts of hydrocarbon exposure on freshwater biodiversity.

Table 3-1 Time weighted average (TWA) concentrations of total polycyclic aromatic compounds (tPACs) from approximately 16 litres of weathered conventional heavy crude oil WAF (water accommodated fraction) poured on June 2nd, 2021. TWA concentrations presented relative to average tadpole sampling event exposure days, presented in ng/L.

Exposure Day	Treatment	Mesocosm ID	Time Weighted Average Concentration at Sampling Time (ng/L)
32	No Oil	5	38.7
32	No Oil	18	69.5
32	No Oil	23	56.7
32	Oil with Plants	3	49.0
32	Oil with Plants	13	66.1
32	Oil with Plants	25	53.3
32	Oil without Plants	9	75.0
32	Oil without Plants	14	57.3
32	Oil without Plants	17	111.4
36	No Oil	5	36.7
36	No Oil	18	67.6
36	No Oil	23	52.9
36	Oil with Plants	3	46.3
36	Oil with Plants	13	62.5
36	Oil with Plants	25	50.3
36	Oil without Plants	9	70.2
36	Oil without Plants	14	54.3
36	Oil without Plants	17	111.6

Table 3-2 Survivorship of tadpoles free-swimming within the mesocosms. Data are expressed as mean and standard deviation.

Treatment	Total Addition ± SD	Known Mortality ± SD	Unknown Fate ± SD	Mesocosm Volume ± SD (L)	Tadpole Density (tadpoles/L)	Survival ± SD (%)
No Oil	79.7 ± 0.6	0.7 ± 0.6	7.3 ± 7.8	1523.89 ± 59.57	0.05	89.9 ± 9.9
Oil with Plants	80.0 ± 0.0	1.0 ± 1.0	27.7 ± 11.6	1504.28 ± 35.42	0.05	64.2 ± 15.6
Oil without Plants	79.3 ± 0.6	1.0 ± 1.0	5.3 ± 1.2	1567.28 ± 71.85	0.05	92.0 ± 2.6

Table 3-3 Mean apical endpoints measured (Gosner stage, snout-vent length, total length, and weight) within each mesocosm at the IISD-Experimental Lakes Area sampled between July 2, 2021 - July 30, 2021, partitioned by sampling event. Data are expressed as mean and standard deviation.

Sampling Event (GS Class)	Treatments	Gosner Stage ± SD	Snout-Vent Length ± SD (mm)	Total Length ± SD (mm)	Weight ± SD (g)
First event (GS 34-36)	No Oil	35.4 ± 0.8	15.7 ± 1.3	37.6 ± 3.8	0.7 ± 0.2
First event (GS 34-36)	Oil with Plants	35.2 ± 0.8	16.5 ± 1.3	38.9 ± 3.0	0.8 ± 0.2
First event (GS 34-36)	Oil without Plants	35.1 ± 0.8	15.9 ± 1.7	38.2 ± 4.4	0.7 ± 0.2
Second event (GS 40-42)	No Oil	40.8 ± 0.7	16.5 ± 1.9	41.7 ± 5.0	0.8 ± 0.3
Second event (GS 40-42)	Oil with Plants	40.7 ± 0.7	17 ± 1.9	42.2 ± 4.6	0.9 ± 0.3
Second event (GS 40-42)	Oil without Plants	40.7 ± 0.7	15.7 ± 2.0	39.6 ± 4.4	0.7 ± 0.2

Table 3-4 Growth Rate from weekly apical endpoints measured (Gosner stage, snout-vent length, total length, and weight) within each mesocosm at the IISD-Experimental Lakes Area sampled between July 2, 2021 - July 30, 2021, with 95% CI. est_k represents the growth rate across treatments, est_b represents the growth rate changes over time and est_r represents the overall growth of the tadpoles.

Treatments	est_k	lower_k	upper_k	est_b	lower_b	upper_b	est_r	lower_r	upper_r
No Oil	0.81	0.73	0.88	3.30	-0.42	7.01	0.17	0.06	0.29
Oil with Plants	0.93	0.68	1.18	2.27	0.11	4.43	0.12	-0.01	0.25
Oil without Plants	0.69	0.57	0.82	3.73	0.62	6.83	0.17	0.07	0.27

Table 3-5 Mean bioaccumulation of total, parent and alkylated polycyclic aromatic compound contents (nanogram/gram of lipid) amongst tadpole tissues. Data from the treatments sampled within each mesocosm at the IISD-Experimental Lakes Area sampled between July 2, 2021 - July 30, 2021partitioned by sampling event. Data are expressed as mean and standard deviation. Concentrations are corrected for lipid content.

Sampling Event (GS Class)	Treatments	PAH Group	Concentration ± SD (ng/g of lipid)
First event (GS 34-36)	No Oil	Total PACs	0.41 ± 0.70
First event (GS 34-36)	Oil with Plants	Total PACs	0.51 ± 0.85
First event (GS 34-36)	Oil without Plants	Total PACs	0.82 ± 0.72
Second Event (GS 40–42)	No Oil	Total PACs	0.16 ± 0.23
Second Event (GS 40–42)	Oil with Plants	Total PACs	0.33 ± 0.29
Second Event (GS 40–42)	Oil without Plants	Total PACs	0.02 ± 0.02
First event (GS 34-36)	No Oil	Parent PACs	0.11 ± 0.16
First event (GS 34-36)	Oil with Plants	Parent PACs	0.08 ± 0.13
First event (GS 34-36)	Oil without Plants	Parent PACs	0.20 ± 0.14
Second Event (GS 40–42)	No Oil	Parent PACs	0.03 ± 0.03
Second Event (GS 40–42)	Oil with Plants	Parent PACs	0.09 ± 0.03
Second Event (GS 40–42)	Oil without Plants	Parent PACs	0.23 ± 0.23
First event (GS 34-36)	No Oil	Alkylated PACs	0.52 ± 0.86
First event (GS 34-36)	Oil with Plants	Alkylated PACs	0.59 ± 0.98
First event (GS 34-36)	Oil without Plants	Alkylated PACs	1.02 ± 0.85
Second Event (GS 40–42)	No Oil	Alkylated PACs	0.19 ± 0.27
Second Event (GS 40–42)	Oil with Plants	Alkylated PACs	0.41 ± 0.26
Second Event (GS 40–42)	Oil without Plants	Alkylated PACs	0.25 ± 0.25

Table 3-6 Mean nearest neighbour distance, percent time spent in inner sector (%) and total distance end-points of weekly behavioural assays conducted on a subset of n=5 tadpole between July 2, 2021 - July 30, 2021 in the mesocosms, partitioned by study day. Data are expressed as mean and standard deviation.

Study Day	Treatments	Total Distance ± SD (cm)	Nearest Neighbour Distance ± SD (cm)	Percent Time in Inner Sector ± SD (%)
7	No Oil	197.5 ± 211.6	7.8 ± 1.3	27.4 ± 18.4
7	Oil with Plants	191.3 ± 135.5	7.7 ± 1.0	39.8 ± 24.0
7	Oil without Plants	105.0 ± 98.6	8.4 ± 2.6	26.0 ± 22.4
15	No Oil	101.6 ± 122.0	7.8 ± 2.0	34.4 ± 31.9
15	Oil with Plants	134.9 ± 127.4	8.1 ± 1.7	30.2 ± 23.8
15	Oil without Plants	156.7 ± 225.1	7.9 ± 3.1	38.3 ± 34.9
21	No Oil	38.2 ± 45.5	9.0 ± 2.9	25.1 ± 27.6
21	Oil with Plants	84.4 ± 78.1	8.6 ± 3.2	27.4 ± 31.5
21	Oil without Plants	78.8 ± 110.8	7.2 ± 1.3	46.1 ± 31.9
28	No Oil	114.9 ± 111.8	7.5 ± 1.3	22.2 ± 27.6
28	Oil with Plants	91.1 ± 86.4	8.3 ± 1.7	19.1 ± 20.4
28	Oil without Plants	206.3 ± 207.0	8.6 ± 2.0	31.2 ± 26.4
28	No Oil - Predator Cue	347.3 ± 192.7	7.5 ± 1.2	21.0 ± 16.7
28	Oil with Plants - Predator Cue	351.6 ± 204.1	7.1 ± 1.5	26.1 ± 25.3
28	Oil without Plants - Predator Cue	252.8 ± 181.3	7.2 ± 1.1	34.1 ± 17.2

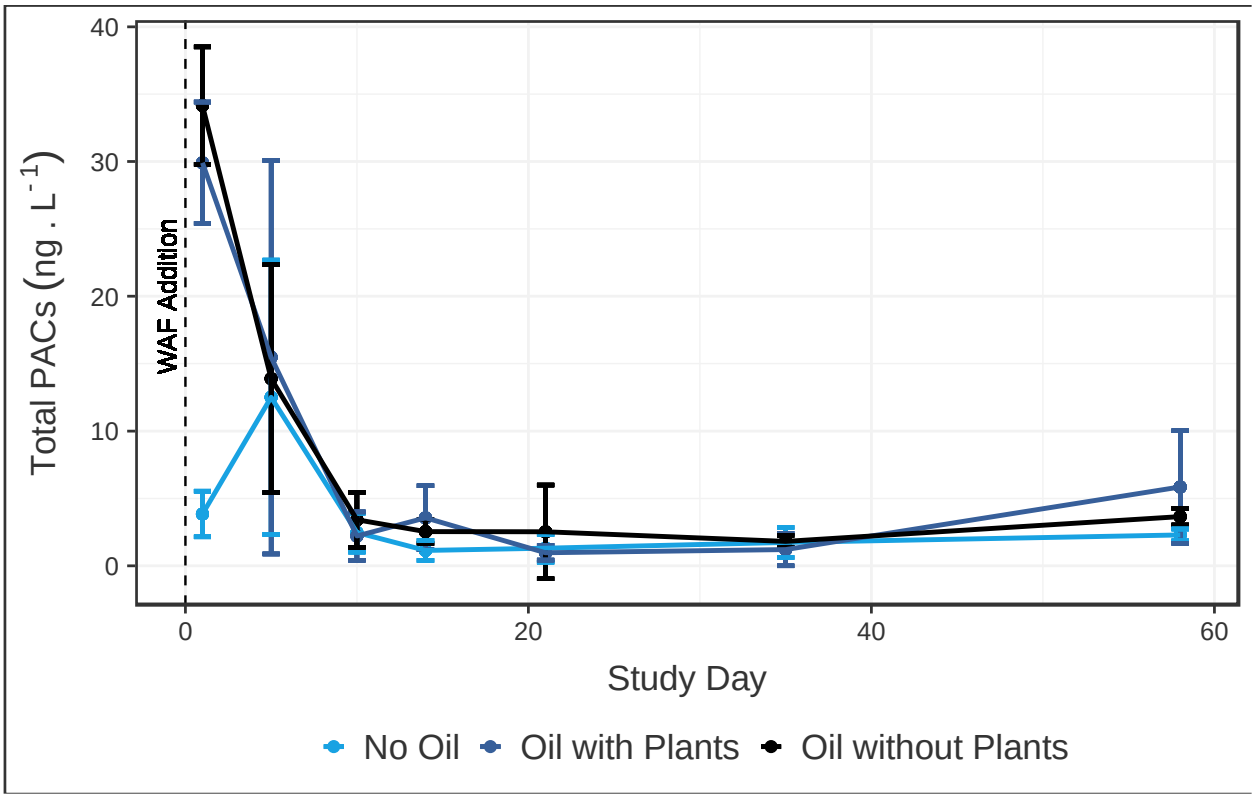


Figure 3-1 16 Priority polycyclic aromatic compound (tPAC) concentrations from approximately 16 liters of weathered conventional heavy crude oil water accommodated fraction (WAF) poured on June 2nd, 2021. Plot represented change in tPAC concentrations over duration of tadpole exposure period within the mesocosms.

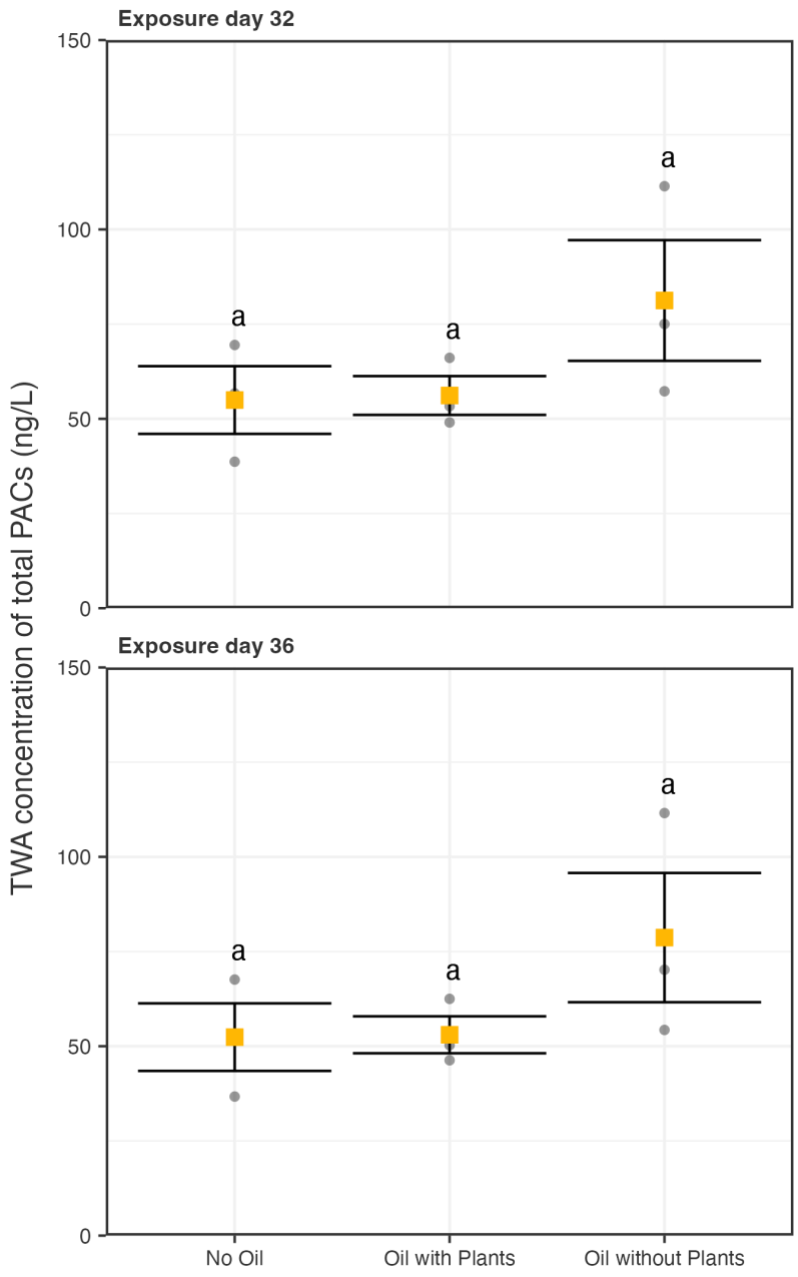


Figure 3-2 Time weighted average (TWA) concentrations of total polycyclic aromatic compounds (tPACs) from approximately 16 litres of weathered conventional heavy crude oil WAF (water accommodated fraction) poured on June 2nd, 2021. TWA concentrations presented relative to average tadpole sampling event exposure days, presented in ng/L. Letters represent statistically significant differences. The multiple comparison test on TWA concentrations revealed no statistically significant differences between the treatments (oil without plants, oil with plants, and no oil) on either exposure day ($p<0.05$).

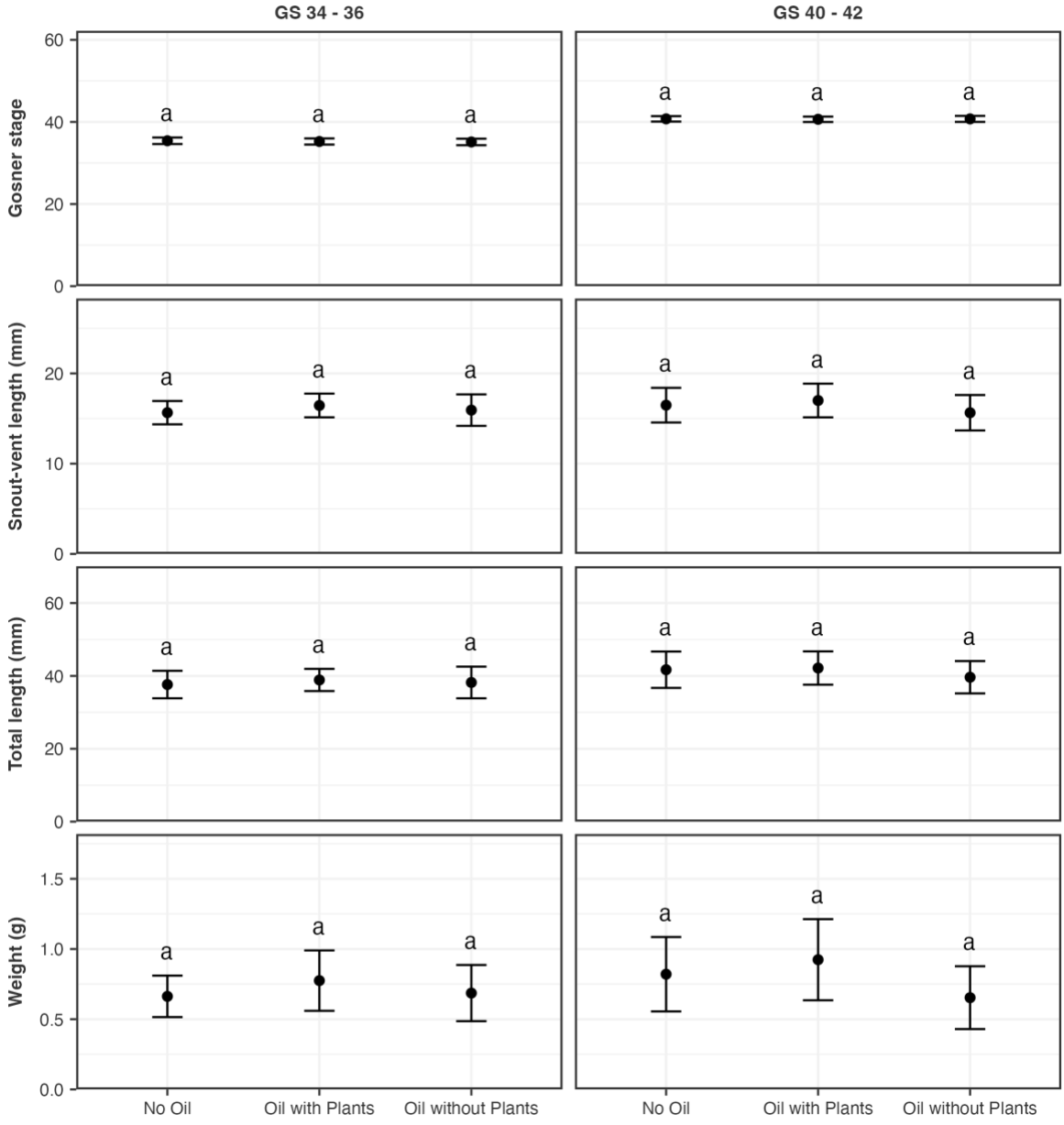


Figure 3-3 Multiple Comparison test on apical endpoints measured (Gosner stage, snout-vent length, total length, and weight) within each mesocosm at the IISD-Experimental Lakes Area sampled between July 2, 2021 - July 30, 2021, partitioned by sampling event. Letters represent statistical significance. The multiple comparisons test data for both sampling events indicate no significant differences in apical endpoints across the treatments; oil without plants, oil with plants, and no oil with plants (control).

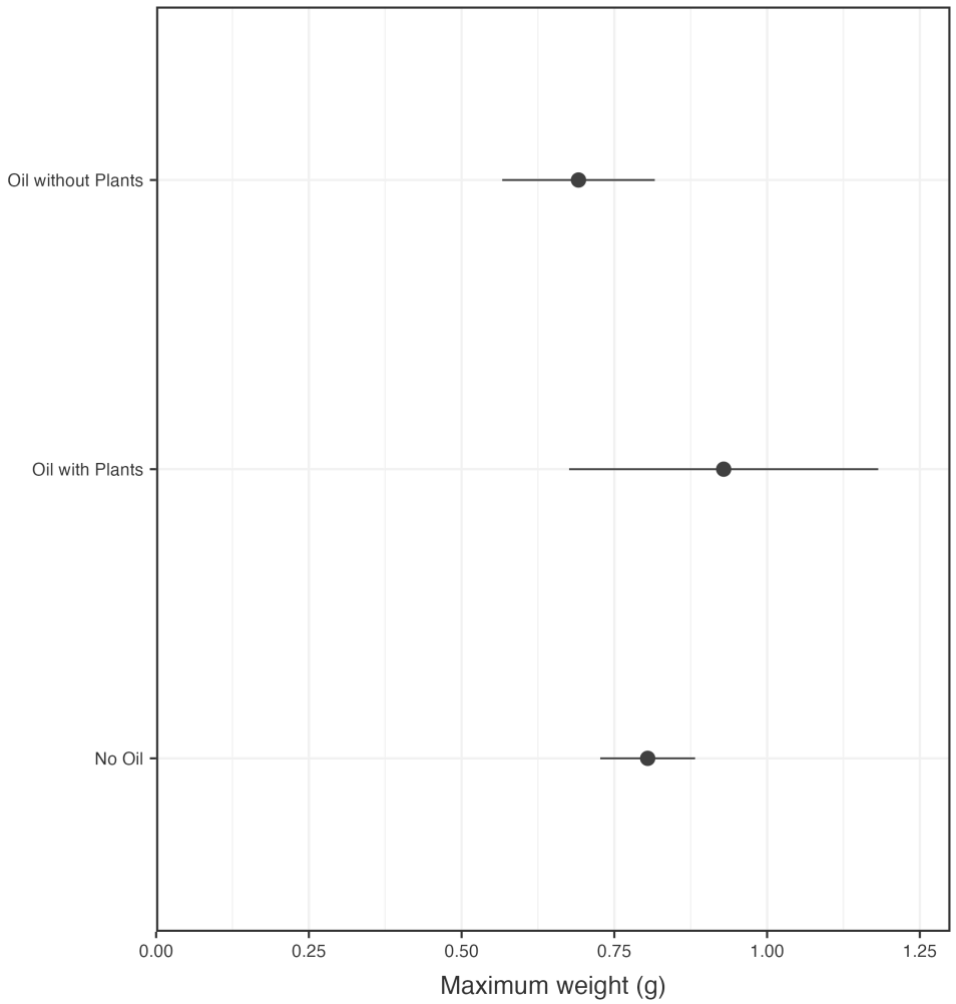


Figure 3-4 Maximum weight (g) based on rate from weekly apical endpoints measured (weight) within each mesocosm at the IISD-Experimental Lakes Area sampled between July 2, 2021 - July 30, 2021.

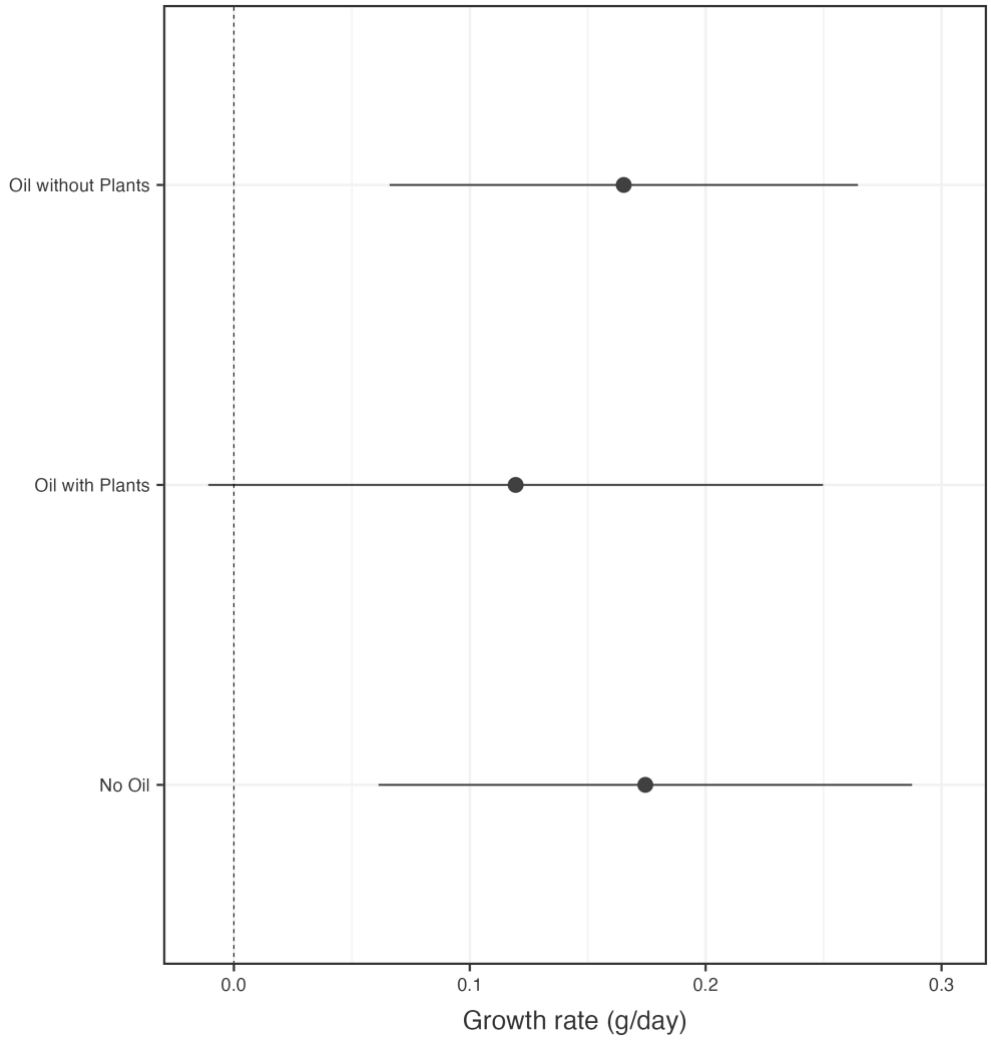


Figure 3-5 Growth Rate measured in grams per day from weekly apical endpoints measured (Gosner stage, snout-vent length, total length, and weight) within each mesocosm at the IISD-Experimental Lakes Area sampled between July 2, 2021 - July 30, 2021.

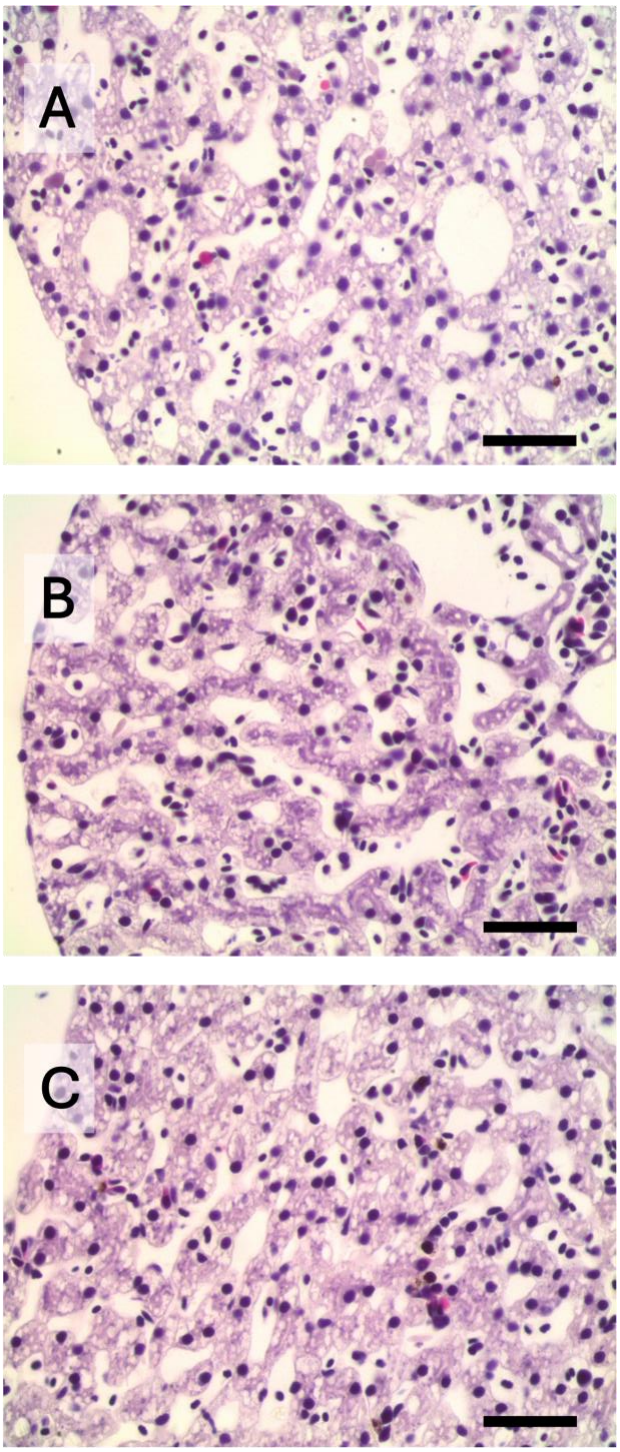


Figure 3-6 Photomicrographs of livers from wood frog (*L. sylvaticus*) tadpoles exposed to various treatments within each *mesocosm* at the IISD-Experimental Lakes Area. (A) No oil, (B) Oil with Plants, (C) Oil without Plants. There was no apparent difference among treatments in the histological appearance of the livers. Scale bar = 50 microns

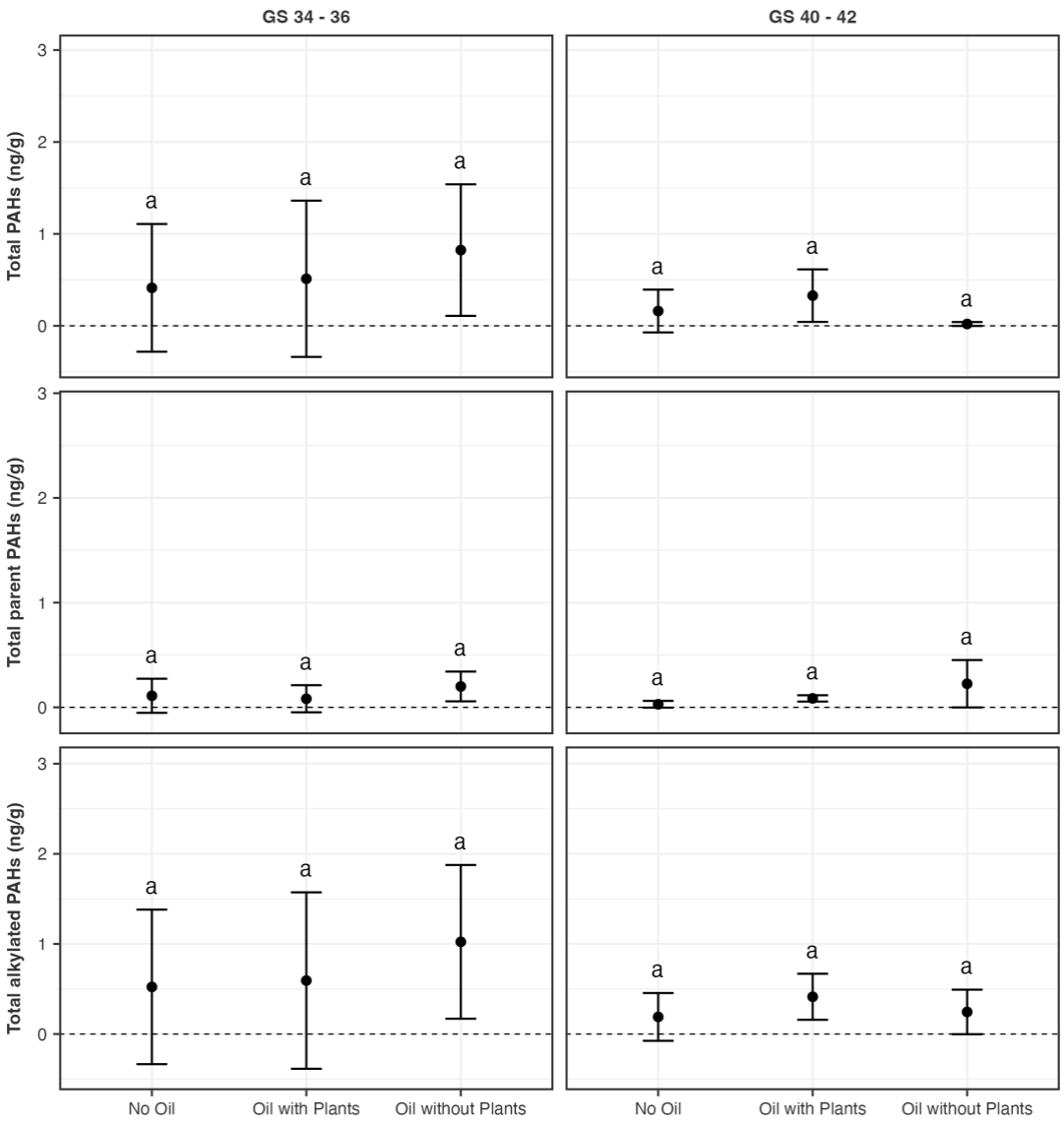


Figure 3-7 Multiple comparisons test of total, parent and alkylated polycyclic aromatic compound contents (ng/gram of lipid) amongst tadpole tissues. Data from the treatments sampled within each mesocosm at the IISD-Experimental Lakes Area sampled between July 2, 2021 - July 30, 2021, partitioned by sampling event. Letters represent statistical significance. Statistical analysis of tPAC bioaccumulation within tadpole tissues of experiment across different treatments and Gosner stages revealed no statistically significant differences in tPACs.

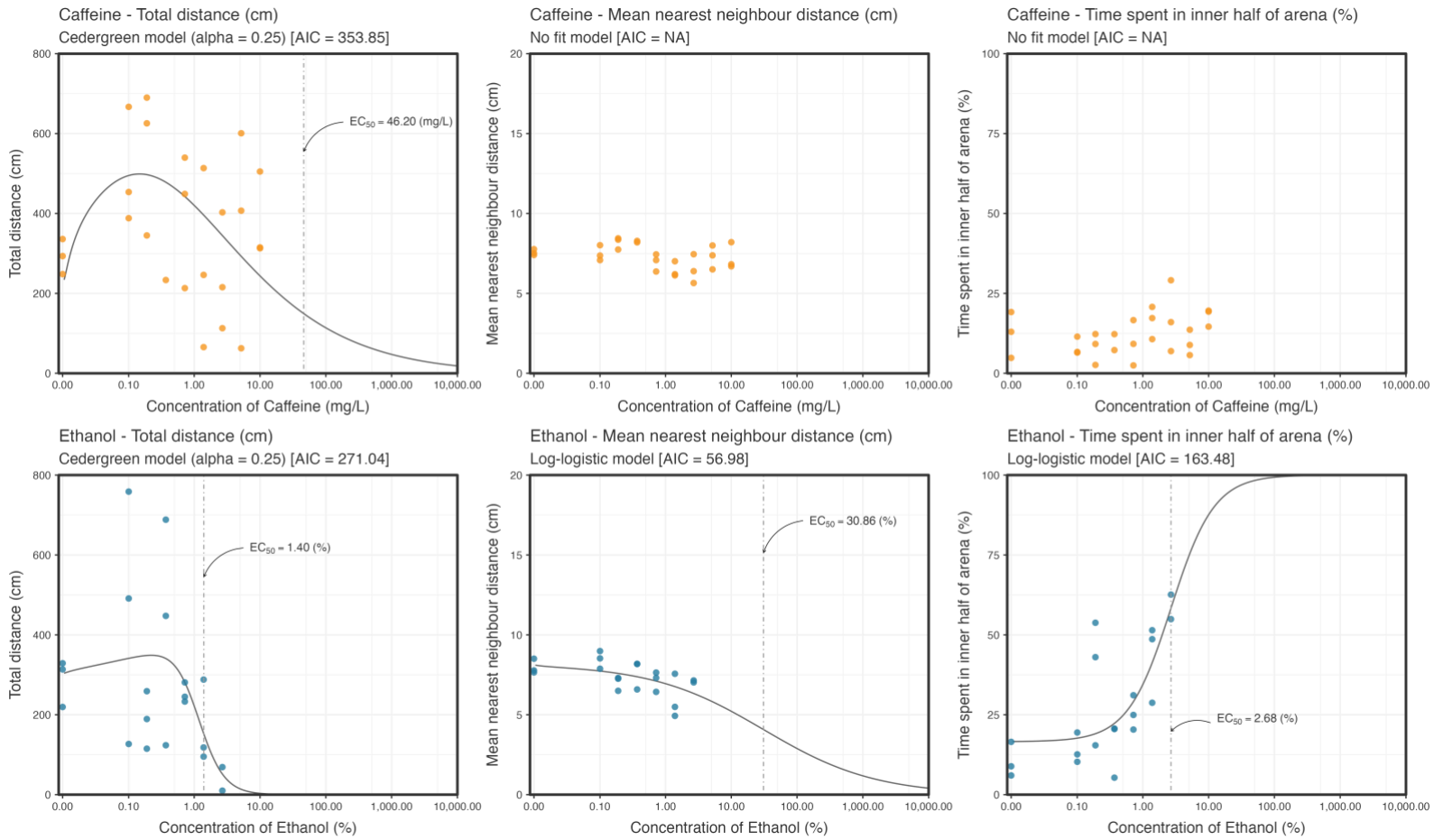


Figure 3-8 Concentration response curves for several behavioural endpoints measured on wood frog exposed to two reference chemicals (caffeine, top, and ethanol, bottom). Panel subtitles indicate the best-fit model used for the regression curve as well as the Akaike information criteria value (AIC) used to determine the best model. EC_{50} values for each compound/endpoint combination are presented where applicable.

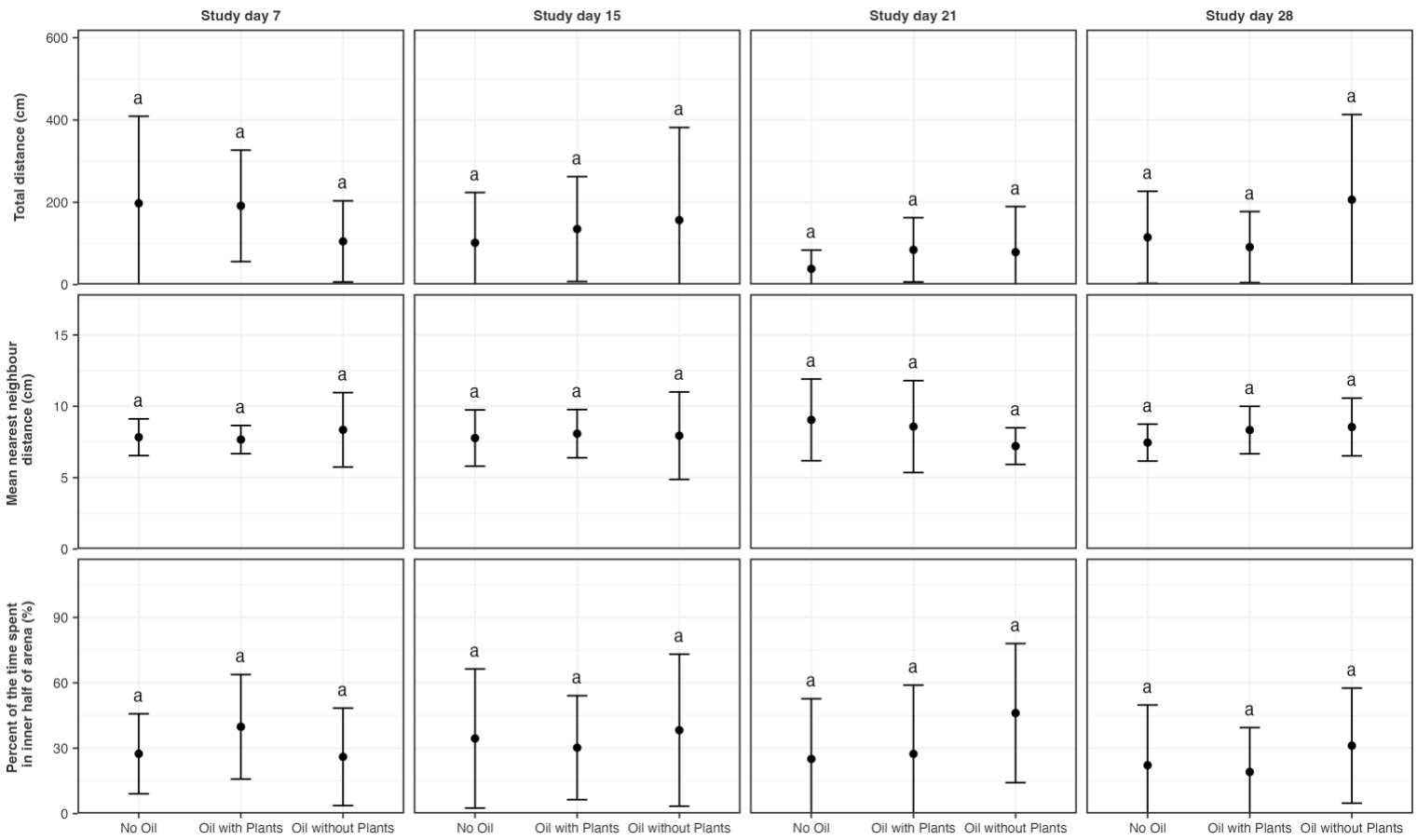


Figure 3-9 Multiple comparisons test of weekly behavioural assays conducted on a subset of n=5 tadpole between July 2, 2021 - July 30, 2021 in the mesocosms, partitioned by study day. Statistical analysis of mean nearest neighbour distance (nnd), percent time spent within inner sector (%) and total distance endpoints measured in the behavioural assays revealed no significant differences between treatments. Letters represent statistical significance.

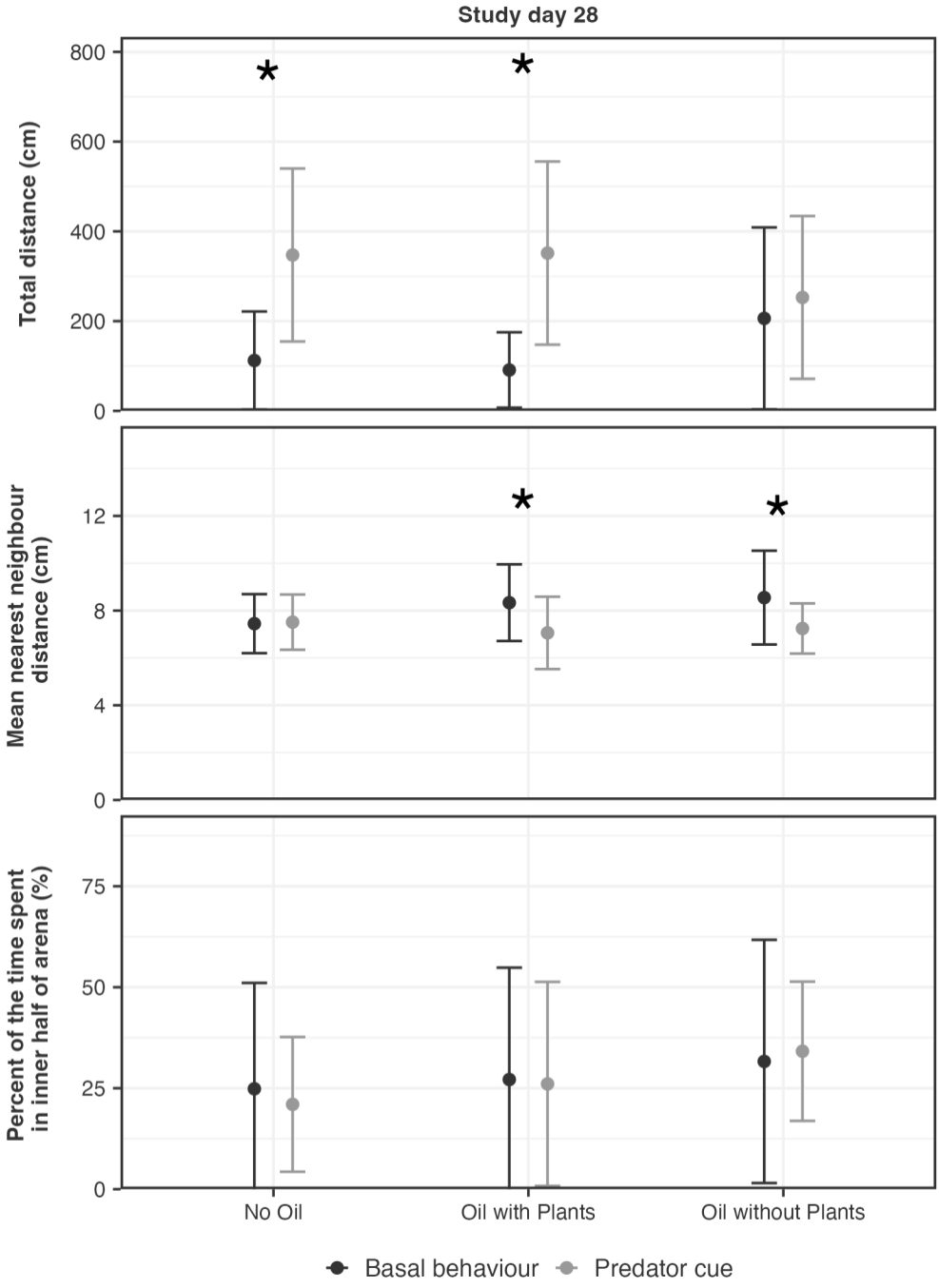


Figure 3-10 Multiple comparisons test of study day 28 behavioural assays conducted on a subset of n=5 tadpoles between July 2, 2021 - July 30, 2021 in the mesocosms, partitioned by study day. Basal behaviour presented in black (left) and behaviour with predator cue addition presented in grey (right). Asterix signifies statistical significance.

SI Table 3-1 List of 44 Polycyclic aromatic compounds (PAC) measured for analysis

PAC Compound	
1,7-Dimethylphenanthrene	C1 Pyrene
1,8-Dimethylphenanthrene	C2 Dibenzothiophene
1-Methylnaphthalene	C2 Fluorene
1-Methylphenanthrene	C2 Naphthalene
2,6-Dimethylphenanthrene	C2 Phenanthrene
2-Methylnaphthalene	C2 Pyrene
2-Methylphenanthrene	C3 Chrysene
3,6-Dimethylphenanthrene	C3 Dibenzothiophene
3-Methylphenanthrene	C3 Naphthalene (Trimethyl)
9/4-Methylphenanthrene	C3 Phenanthrene
Acenapthene	C4 Naphthalene (Tetramethyl)
Acenapthylene	C4 Phenanthrene
Anthracene	Chrysene
Benz[a]anthracene	Dibenzo[a,h]anthracene
Benzo[a]pyrene	Dibenzothiophene
Benzo[b]fluoranthene	Fluoranthene
Benzo[g,h,i]perylene	Fluorene
Benzo[k]fluoranthene	Indeno[1,2,3-c,d]pyrene
C1 Benzo[a]pyrene	Naphthalene
C1 Chrysene	Phenanthrene
C1 Dibenzothiophene	Pyrene
C1 Fluorene	Retene

SI Table 3-2 Tadpole histology samples collected in 2021.

Mesocosm ID	Treatment	Number of Tadpoles
5	No Oil	6
18	No Oil	6
23	No Oil	6
9	Oil without plants	7
14	Oil without plants	6
17	Oil without plants	6
3	Oil with plants	1
13	Oil with plants	6
25	Oil with plants	5

SI Table 3-3 Pooled sample mass for total polycyclic aromatic compound (tPAC) bioaccumulation analysis of tadpole tissues. To achieve sufficient total sample mass after removal of GI tract, for the quantification of tPAC bioaccumulation in tadpole tissues, tadpoles from the mesocosms were pooled in groups of four (n=4) from the same treatments and same Gosner stage classes

Pool ID (n-4)	Treatment	GS Class	Total Sample Mass (grams)
pool_01	Oil with Plants	GS 34 - 36	1.05
pool_02	Oil with Plants	GS 40 - 42	1.18
pool_03	No Oil	GS 34 - 36	1.05
pool_04	No Oil	GS 40 - 42	1.42
pool_05	Oil without Plants	GS 34 - 36	0.95
pool_06	Oil without Plants	GS 40 - 42	1.19
pool_07	Oil with Plants	GS 34 - 36	1.06
pool_08	Oil with Plants	GS 40 - 42	1.86
pool_09	Oil without Plants	GS 34 - 36	1.13
pool_10	Oil without Plants	GS 40 - 42	1.40
pool_11	Oil without Plants	GS 34 - 36	0.71
pool_12	Oil without Plants	GS 40 - 42	0.84
pool_13	No Oil	GS 34 - 36	1.04
pool_14	No Oil	GS 40 - 42	1.38
pool_15	No Oil	GS 34 - 36	0.83
pool_16	No Oil	GS 40 - 42	1.47
pool_17	Oil with Plants	GS 34 - 36	1.04
pool_18	Oil with Plants	GS 40 - 42	1.55

SI Table 3-4 Behavioural Assay compounds (Ethanol (%) and Caffeine (mg/L)) and the dosages

Compound	Treatment	Dose
Caffeine	Control	0 mg/L
Caffeine	Treatment 1	0.1 mg/L
Caffeine	Treatment 2	0.19 mg/L
Caffeine	Treatment 3	0.37 mg/L
Caffeine	Treatment 4	0.72 mg/L
Caffeine	Treatment 5	1.39 mg/L
Caffeine	Treatment 6	2.68 mg/L
Caffeine	Treatment 7	5.18 mg/L
Caffeine	Treatment 8	10 mg/L
Ethanol	Control	0 %
Ethanol	Treatment 1	0.001 %
Ethanol	Treatment 2	0.0019 %
Ethanol	Treatment 3	0.0037 %
Ethanol	Treatment 4	0.0072 %
Ethanol	Treatment 5	0.0139 %
Ethanol	Treatment 6	0.0268 %

SI Table 3-5 Mean volumes of water in mesocosms partitioned by mesocosm identifications from May 30, 2021 - July 30, 2021. Volumes (L) averaged, and SD presented.

Mesocosm ID	Treatment	Mean Volume ± SD (L)
3	Oil with Plants	1465.3 ± 134.4
5	No Oil	1485.3 ± 145.9
9	Oil without Plants	1551.5 ± 128.4
13	Oil with Plants	1513.1 ± 131.1
14	Oil without Plants	1504.7 ± 139.5
17	Oil without Plants	1645.7 ± 122.7
18	No Oil	1493.9 ± 137.6
23	No Oil	1592.5 ± 145.8
25	Oil with Plants	1534.5 ± 157.2

SI Table 3-6 Mean exposure duration of tadpoles within mesocosms partitioned by sampling event from July 2, 2021 - July 30, 2021. Exposure duration averaged and SD presented.

Sampling Event (GS Class)	Mean Sampled Study Day ± SD
First event (GS 34-36)	21.6 ± 1.9
Second event (GS 40-42)	28.00 ± 0.0

SI Table 3-7 DO (mg/L), DO (%), pH, Specific Conductivity (mS/cm), Temperature (°C) within the (n=7) rearing tanks at the IISD-Experimental Lakes Area from May 4th - June 18th, 2021. Parameters were measured daily with a YSI handheld probe. Data are expressed as mean and standard deviation.

Parameter	Mean Parameter ± SD
DO (mg/L)	10.5 ± 6.0
DO (%)	92.9 ± 16.4
pH	6.3 ± 0.3
Specific Conductivity (mS/cm)	57.1 ± 48.5
Temperature (°C)	12.5 ± 5.2

SI Table 3-8 Mean temperature measured by HOBO loggers present in the mesocosms from between May 30, 2021 - July 30, 2021, partitioned by treatment. Temperatures were recorded with a HOBO temperature logger, every 5 minutes. Data are expressed as mean and standard deviation.

Treatments	Temperature ± SD (°C)
No Oil	21.1 ± 4.1
Oil with Plants	21.5 ± 4.0
Oil without Plants	21.9 ± 4.1

SI Table 3-9 ANOVA of data from HOBO loggers present in the mesocosms from between May 30, 2021 - July 30, 2021. Temperatures (°C) were recorded with a HOBO temperature logger, every 5 minutes.

Term	df	SumSQ	MeanSQ	Statistic	P-Value
Treatments	2	0.38	0.19	8.29	0.24
Residuals	1	0.02	0.02	NA	NA

SI Table 3-10 Multiple Comparisons of data from HOBO loggers present in the mesocosms from between May 30, 2021 - July 30, 2021. Temperatures were recorded with a HOBO temperature logger, every 5 minutes. Hobo loggers were not present within every treatment replicate, and therefore, P-value cannot be evaluated

Contrast	Estimate	SE	df	T-Ratio	P-Value
No Oil - Oil with Plants	-0.41	0.18	1	-2.24	NaN
No Oil - Oil without Plants	-0.87	0.21	1	-4.07	NaN
Oil with Plants - Oil without Plants	-0.45	0.18	1	-2.46	NaN

SI Table 3-11 Mean DO (mg/L), DO (%), pH, Specific Conductivity (mS/cm), Temperature (°C) partitioned by mesocosm treatments. Parameters were measured with a YSI handheld probe from May 30, 2021 - July 30, 2021. Data are expressed as mean and standard deviation.

Treatment	Dissolved Oxygen ± SD (%)	Dissolved Oxygen ± SD (mg/L)	pH ± SD	Specific Conductivity ± SD (mS/cm)	Temperature ± SD (°C)
No Oil	90.2 ± 10.6	8.0 ± 1.0	6.0 ± 0.4	22.0 ± 2.4	21.1 ± 2.4
Oil with Plants	94.1 ± 9.6	8.3 ± 0.8	6.4 ± 2.7	20.5 ± 3.6	21.4 ± 2.4
Oil without Plants	97.3 ± 7.8	11.3 ± 14.1	6.0 ± 0.2	21.6 ± 2.1	21.5 ± 2.3

SI Table 3-12 ANOVA of DO (mg/L), DO (%), pH, Specific Conductivity (mS/cm), Temperature (°C) partitioned by mesocosm treatments. Parameters were measured with a YSI handheld probe from May 30, 2021 - July 30, 2021.

Parameter	Term	df	SumSQ	MeanSQ	Statistic	P-Value
Dissolved Oxygen (mg/L)	Treatments	2	19.23	9.61	1.39	0.32
Dissolved Oxygen (mg/L)	Residuals	6	41.36	6.89	NA	NA
Dissolved Oxygen (%)	Treatments	2	75.48	37.74	19.28	0
Dissolved Oxygen (%)	Residuals	6	11.74	1.96	NA	NA
pH	Treatments	2	0.37	0.18	0.68	0.54
pH	Residuals	6	1.64	0.27	NA	NA
Specific Conductivity (mS/cm)	Treatments	2	3.61	1.81	2.72	0.14
Specific Conductivity (mS/cm)	Residuals	6	3.98	0.66	NA	NA
Temperature (°C)	Treatments	2	0.28	0.14	0.8	0.49
Temperature (°C)	Residuals	6	1.05	0.17	NA	NA

SI Table 3-13 Multiple Comparisons of DO (mg/L), DO (%), pH, Specific Conductivity (mS/cm), Temperature (°C) partitioned by mesocosm treatments. Parameters were measured with a YSI handheld probe from May 30, 2021 - July 30, 2021.

Parameter	Contrast	Estimate	SE	df	T-Ratio	P-Value
Dissolved Oxygen (mg/L)	No Oil - Oil with Plants	-0.24	2.14	6	-0.11	0.99
Dissolved Oxygen (mg/L)	No Oil - Oil without Plants	-3.21	2.14	6	-1.5	0.36
Dissolved Oxygen (mg/L)	Oil with Plants - Oil without Plants	-2.98	2.14	6	-1.39	0.4
Dissolved Oxygen (%)	No Oil - Oil with Plants	-3.84	1.14	6	-3.36	0.03
Dissolved Oxygen (%)	No Oil - Oil without Plants	-7.09	1.14	6	-6.2	0
Dissolved Oxygen (%)	Oil with Plants - Oil without Plants	-3.24	1.14	6	-2.84	0.07
pH	No Oil - Oil with Plants	-0.41	0.43	6	-0.96	0.63
pH	No Oil - Oil without Plants	0.04	0.43	6	0.09	1
pH	Oil with Plants - Oil without Plants	0.45	0.43	6	1.05	0.58
Specific Conductivity (mS/cm)	No Oil - Oil with Plants	1.5	0.66	6	2.26	0.14
Specific Conductivity (mS/cm)	No Oil - Oil without Plants	0.4	0.66	6	0.61	0.82
Specific Conductivity (mS/cm)	Oil with Plants - Oil without Plants	-1.09	0.66	6	-1.65	0.3
Temperature (°C)	No Oil - Oil with Plants	-0.36	0.34	6	-1.05	0.58
Temperature (°C)	No Oil - Oil without Plants	-0.39	0.34	6	-1.14	0.53
Temperature (°C)	Oil with Plants - Oil without Plants	-0.03	0.34	6	-0.09	1

SI Table 3-14 Mean water nutrient parameters within the mesocosms. Nutrient parameter conditions relevant to tadpole include alkalinity (Alk), ammonia (NH3), calcium (Ca), chlorine (Cl_2), dis-solved inorganic carbon (DIC), dissolved organic carbon (DOC), particulate carbon (PC), particulate phosphorus (PP), particulate nitrogen (PN), total dissolved phosphorus TDP), total dissolved nitrogen (TDN), total dissolved solids (TDS), total suspended solids (TSS). Water nutrient parameters were sampled between May 30, 2021 - July 30, 2021 and analyzed by the IISD-ELA Chemistry Laboratory at the IISD-Experimental Lakes Area. Data are expressed as mean and standard deviation.

Parameter	Treatment	Mean Parameter ± SD
Alkalinity (Alk)	No Oil	96.93 µEq/L ± 10.21
Alkalinity (Alk)	Oil with Plants	108.09 µEq/L ± 28.57
Alkalinity (Alk)	Oil without Plants	108.26 µEq/L ± 9.02
Ammonia (NH_3)	No Oil	74.43 µg/L ± 103.10
Ammonia (NH_3)	Oil with Plants	45.29 µg/L ± 71.95
Ammonia (NH_3)	Oil without Plants	90.21 µg/L ± 93.12
Calcium (Ca)	No Oil	1.48 mg/L ± 0.15
Calcium (Ca)	Oil with Plants	1.63 mg/L ± 0.18
Calcium (Ca)	Oil without Plants	1.65 mg/L ± 0.18
Chlorine (Cl_2)	No Oil	0.25 mg/L ± 0.05
Chlorine (Cl_2)	Oil with Plants	0.22 mg/L ± 0.06
Chlorine (Cl_2)	Oil without Plants	0.25 mg/L ± 0.06
Dissolved Inorganic Carbon (DIC)	No Oil	107.07 µmol/L ± 17.38
Dissolved Inorganic Carbon (DIC)	Oil with Plants	108.47 µmol/L ± 17.97
Dissolved Inorganic Carbon (DIC)	Oil without Plants	99.54 µmol/L ± 11.83
Dissolved Organic Carbon (DOC)	No Oil	765.64 µmol/L ± 113.40
Dissolved Organic Carbon (DOC)	Oil with Plants	795.00 µmol/L ± 131.38
Dissolved Organic Carbon (DOC)	Oil without Plants	786.21 µmol/L ± 124.86
Particulate Carbon (PC)	No Oil	392.60 µg/L ± 84.61
Particulate Carbon (PC)	Oil with Plants	437.74 µg/L ± 99.65
Particulate Carbon (PC)	Oil without Plants	404.99 µg/L ± 134.68
Particulate Nitrogen (PN)	No Oil	46.34 µg/L ± 11.27

Particulate Nitrogen (PN)	Oil with Plants	49.74 µg/L ± 15.89
Particulate Nitrogen (PN)	Oil without Plants	49.24 µg/L ± 16.09
Particulate Phosphorus (PP)	No Oil	4.29 µg/L ± 1.20
Particulate Phosphorus (PP)	Oil with Plants	5.14 µg/L ± 1.10
Particulate Phosphorus (PP)	Oil without Plants	3.64 µg/L ± 1.34
Total Dissolved Nitrogen (TDN)	No Oil	963.36 µg/L ± 223.23
Total Dissolved Nitrogen (TDN)	Oil with Plants	943.29 µg/L ± 190.08
Total Dissolved Nitrogen (TDN)	Oil without Plants	1008.86 µg/L ± 192.03
Total Dissolved Phosphorus (TDP)	No Oil	5.82 µg/L ± 1.02
Total Dissolved Phosphorus (TDP)	Oil with Plants	5.71 µg/L ± 1.67
Total Dissolved Phosphorus (TDP)	Oil without Plants	4.36 µg/L ± 1.07
Total Dissolved Solids (TDS)	No Oil	44.50 mg/L ± 20.11
Total Dissolved Solids (TDS)	Oil with Plants	34.79 mg/L ± 22.86
Total Dissolved Solids (TDS)	Oil without Plants	59.43 mg/L ± 56.22
Total Suspended Solids (TSS)	No Oil	1.02 mg/L ± 0.85
Total Suspended Solids (TSS)	Oil with Plants	1.55 mg/L ± 0.97
Total Suspended Solids (TSS)	Oil without Plants	1.09 mg/L ± 1.17

SI Table 3-15 ANOVA of water nutrient parameters within the mesocosms. Nutrient parameter conditions relevant to tadpole include alkalinity (Alk), ammonia (NH3), calcium (Ca), chlorine (Cl_2), dissolved inorganic carbon (DIC), dissolved organic carbon (DOC), particulate carbon (PC), particulate phosphorus (PP), particulate nitrogen (PN), total dissolved phosphorus TDP), total dissolved nitrogen (TDN), total dissolved solids (TDS), total suspended solids (TSS). Water nutrient parameters were sampled between May 30, 2021 - July 30, 2021 and analyzed by the IISD-ELA Chemistry Laboratory at the IISD-Experimental Lakes Area.

Parameter	Term	df	SumSQ	MeanSQ	Statistic	P-Value
Alkalinity (Alk)	Treatments	2	126.8	63.4	0.58	0.59
Alkalinity (Alk)	Residuals	6	655.1	109.18	NA	NA
Ammonia (NH_3)	Treatments	2	3443.05	1721.52	5.35	0.05
Ammonia (NH_3)	Residuals	6	1929.8	321.63	NA	NA
Calcium (Ca)	Treatments	2	0.01	0.01	0.21	0.81
Calcium (Ca)	Residuals	6	0.17	0.03	NA	NA
Chlorine (Cl_2)	Treatments	2	0.01	0	1.37	0.32
Chlorine (Cl_2)	Residuals	6	0.02	0	NA	NA
Dissolved Inorganic Carbon (DIC)	Treatments	2	122.83	61.41	2.09	0.21
Dissolved Inorganic Carbon (DIC)	Residuals	6	176.68	29.45	NA	NA
Dissolved Organic Carbon (DOC)	Treatments	2	2830.89	1415.44	0.53	0.61
Dissolved Organic Carbon (DOC)	Residuals	6	15918.2	2653.03	NA	NA
Particulate Carbon (PC)	Treatments	2	5598.9	2799.45	0.4	0.69
Particulate Carbon (PC)	Residuals	6	42448.09	7074.68	NA	NA
Particulate Nitrogen (PN)	Treatments	2	202.62	101.31	0.59	0.58
Particulate Nitrogen (PN)	Residuals	6	1022.6	170.43	NA	NA
Particulate Phosphorus (PP)	Treatments	2	2.33	1.16	2.32	0.18
Particulate Phosphorus (PP)	Residuals	6	3.01	0.5	NA	NA
Total Dissolved Nitrogen (TDN)	Treatments	2	16897.52	8448.76	2.61	0.15
Total Dissolved Nitrogen (TDN)	Residuals	6	19432.01	3238.67	NA	NA
Total Dissolved Phosphorus (TDP)	Treatments	2	2.42	1.21	3.39	0.1

Total Dissolved Phosphorus (TDP)	Residuals	6	2.14	0.36	NA	NA
Total Dissolved Solids (TDS)	Treatments	2	769.11	384.56	3.56	0.1
Total Dissolved Solids (TDS)	Residuals	6	647.69	107.95	NA	NA
Total Suspended Solids (TSS)	Treatments	2	0.17	0.08	1.22	0.36
Total Suspended Solids (TSS)	Residuals	6	0.41	0.07	NA	NA

SI Table 3-16 Multiple Comparisons of water nutrient parameters within the mesocosms. Nutrient parameter conditions relevant to tadpole include alkalinity (Alk), ammonia (NH3), calcium (Ca), chlorine (Cl_2), dissolved inorganic carbon (DIC), dissolved organic carbon (DOC), particulate carbon (PC), particulate phosphorus (PP), particulate nitrogen (PN), total dissolved phosphorus TDP), total dissolved nitrogen (TDN), total dissolved solids (TDS), total suspended solids (TSS). Water nutrient parameters were sampled between May 30, 2021 - July 30, 2021, and analyzed by the IISD-ELA Chemistry Laboratory at the IISD-Experimental Lakes Area.

Parameter	Contrast	Estimate	SE	df	T-Ratio	P-Value
Alkalinity	No Oil - Oil with Plants	-4.83	8.53	6	-0.57	0.84
Alkalinity	No Oil - Oil without Plants	-9.19	8.53	6	-1.08	0.56
Alkalinity	Oil with Plants - Oil without Plants	-4.36	8.53	6	-0.51	0.87
Ammonia	No Oil - Oil with Plants	17.86	14.64	6	1.22	0.49
Ammonia	No Oil - Oil without Plants	-29.57	14.64	6	-2.02	0.19
Ammonia	Oil with Plants - Oil without Plants	-47.43	14.64	6	-3.24	0.04
Calcium	No Oil - Oil with Plants	0.01	0.14	6	0.1	0.99
Calcium	No Oil - Oil without Plants	-0.07	0.14	6	-0.51	0.87
Calcium	Oil with Plants - Oil without Plants	-0.08	0.14	6	-0.61	0.82
Chlorine	No Oil - Oil with Plants	0.07	0.04	6	1.5	0.35
Chlorine	No Oil - Oil without Plants	0.06	0.04	6	1.35	0.42
Chlorine	Oil with Plants - Oil without Plants	-0.01	0.04	6	-0.15	0.99
Dissolved Inorganic Carbon	No Oil - Oil with Plants	-0.01	4.43	6	0	1
Dissolved Inorganic Carbon	No Oil - Oil without Plants	7.83	4.43	6	1.77	0.26
Dissolved Inorganic Carbon	Oil with Plants - Oil without Plants	7.84	4.43	6	1.77	0.26
Dissolved Organic Carbon	No Oil - Oil with Plants	16.14	42.06	6	0.38	0.92
Dissolved Organic Carbon	No Oil - Oil without Plants	-26.86	42.06	6	-0.64	0.81
Dissolved Organic Carbon	Oil with Plants - Oil without Plants	-43	42.06	6	-1.02	0.59
Particulate Carbon	No Oil - Oil with Plants	37.29	68.68	6	0.54	0.85
Particulate Carbon	No Oil - Oil without Plants	60.56	68.68	6	0.88	0.67
Particulate Carbon	Oil with Plants - Oil without Plants	23.27	68.68	6	0.34	0.94
Particulate Nitrogen	No Oil - Oil with Plants	9.84	10.66	6	0.92	0.65
Particulate Nitrogen	No Oil - Oil without Plants	10.27	10.66	6	0.96	0.62

Particulate Nitrogen	Oil with Plants - Oil without Plants	0.43	10.66	6	0.04	1
Particulate Phosphorus	No Oil - Oil with Plants	-0.14	0.58	6	-0.25	0.97
Particulate Phosphorus	No Oil - Oil without Plants	1	0.58	6	1.73	0.27
Particulate Phosphorus	Oil with Plants - Oil without Plants	1.14	0.58	6	1.98	0.2
Total Dissolved Nitrogen	No Oil - Oil with Plants	49.24	46.47	6	1.06	0.57
Total Dissolved Nitrogen	No Oil - Oil without Plants	-56.81	46.47	6	-1.22	0.48
Total Dissolved Nitrogen	Oil with Plants - Oil without Plants	-106.05	46.47	6	-2.28	0.13
Total Dissolved Phosphorus	No Oil - Oil with Plants	0.57	0.49	6	1.16	0.51
Total Dissolved Phosphorus	No Oil - Oil without Plants	1.27	0.49	6	2.6	0.09
Total Dissolved Phosphorus	Oil with Plants - Oil without Plants	0.7	0.49	6	1.44	0.38
Total Dissolved Solids	No Oil - Oil with Plants	8.53	8.48	6	1.01	0.6
Total Dissolved Solids	No Oil - Oil without Plants	-13.9	8.48	6	-1.64	0.3
Total Dissolved Solids	Oil with Plants - Oil without Plants	-22.43	8.48	6	-2.64	0.09
Total Suspended Solids	No Oil - Oil with Plants	-0.32	0.21	6	-1.48	0.37
Total Suspended Solids	No Oil - Oil without Plants	-0.06	0.21	6	-0.3	0.95
Total Suspended Solids	Oil with Plants - Oil without Plants	0.25	0.21	6	1.18	0.51

SI Table 3-17 Multiple Comparison test on time weighted average (TWA) concentrations of total polycyclic aromatic compounds (tPACs) from approximately 16 litres of weathered conventional heavy crude oil WAF (water accommodated fraction) poured on June 2nd, 2021. The multiple comparisons test data for both exposure day indicate no significant differences in TWA concentrations of tPACs across the treatments; oil without plants, oil with plants, and no oil with plants (control).

Exposure Day	Contrast	Estimate	SE	DF	T.Ratio	P.Value
32	No Oil - Oil with Plants	-1.20	15.49	6	-0.08	1.00
32	No Oil - Oil without Plants	-26.28	15.49	6	-1.70	0.28
32	Oil with Plants - Oil without Plants	-25.08	15.49	6	-1.62	0.31
36	No Oil - Oil with Plants	-0.62	16.22	6	-0.04	1.00
36	No Oil - Oil without Plants	-26.28	16.22	6	-1.62	0.31
36	Oil with Plants - Oil without Plants	-25.66	16.22	6	-1.58	0.32

SI Table 3-18 Mixed Effect Model on apical endpoints measured (Gosner stage, snout-vent length, total length, and weight) within each mesocosm at the IISD-Experimental Lakes Area sampled between July 2, 2021 - July 30, 2021, partitioned by sampling event. F-statistics indicate minor differences in variability among the different apical endpoints measured in each GS class, suggesting that there is no strong evidence of significant differences in the measured apical endpoints across the treatments.

Sampling Event (GS Class)	Endpoint	npar	sumsq	meansq	Statistic
First event (GS 34-36)	Weight (g)	2	0.051	0.025	0.769
First event (GS 34-36)	Gosner Stage	2	0.612	0.306	0.517
First event (GS 34-36)	Snout Vent Length (mm)	2	2.849	1.425	0.659
First event (GS 34-36)	Total Length (mm)	2	10.551	5.275	0.358
Second Event (GS 40–42)	Weight (g)	2	0.263	0.131	2.394
Second Event (GS 40–42)	Gosner Stage	2	0.43	0.215	0.481
Second Event (GS 40–42)	Snout Vent Length (mm)	2	12.584	6.292	1.95
Second Event (GS 40–42)	Total Length (mm)	2	35.288	17.644	0.897

SI Table 3-19 Multiple Comparison test on apical endpoints measured (Gosner stage, snout-vent length, total length, and weight) within each mesocosm at the IISD-Experimental Lakes Area sampled between July 2, 2021 - July 30, 2021, partitioned by sampling event. The multiple comparisons test data for both sampling events indicate no significant differences in apical endpoints across the treatments; oil without plants, oil with plants, and no oil with plants (control).

Sampling Event (GS Class)	Endpoint	Contrast	Estimate	SE	DF	T.Ratio	P.Value
First event (GS 34-36)	Weight (g)	No Oil - Oil with Plants	-0.11	0.09	7.30	-1.20	0.49
First event (GS 34-36)	Weight (g)	No Oil - Oil without Plants	-0.03	0.08	5.66	-0.38	0.93
First event (GS 34-36)	Weight (g)	Oil with Plants - Oil without Plants	0.07	0.08	5.67	0.90	0.66
First event (GS 34-36)	Gosner Stage	No Oil - Oil with Plants	0.17	0.31	8.30	0.55	0.85
First event (GS 34-36)	Gosner Stage	No Oil - Oil without Plants	0.28	0.28	5.66	1.01	0.60
First event (GS 34-36)	Gosner Stage	Oil with Plants - Oil without Plants	0.11	0.28	5.71	0.41	0.91
First event (GS 34-36)	Snout Vent Length (mm)	No Oil - Oil with Plants	-0.75	0.66	7.57	-1.14	0.52
First event (GS 34-36)	Snout Vent Length (mm)	No Oil - Oil without Plants	-0.31	0.61	5.65	-0.51	0.87
First event (GS 34-36)	Snout Vent Length (mm)	Oil with Plants - Oil without Plants	0.44	0.61	5.66	0.71	0.77
First event (GS 34-36)	Total Length (mm)	No Oil - Oil with Plants	-1.27	1.52	8.18	-0.84	0.69
First event (GS 34-36)	Total Length (mm)	No Oil - Oil without Plants	-0.70	1.39	5.57	-0.51	0.87
First event (GS 34-36)	Total Length (mm)	Oil with Plants - Oil without Plants	0.57	1.36	5.46	0.42	0.91
Second Event (GS 40 – 42)	Weight (g)	No Oil - Oil with Plants	-0.10	0.13	6.08	-0.78	0.73
Second Event (GS 40 – 42)	Weight (g)	No Oil - Oil without Plants	0.18	0.13	5.71	1.41	0.40
Second Event (GS 40 – 42)	Weight (g)	Oil with Plants - Oil without Plants	0.29	0.13	6.21	2.15	0.16
Second Event (GS 40 – 42)	Gosner Stage	No Oil - Oil with Plants	0.17	0.20	6.18	0.84	0.69
Second Event (GS 40 – 42)	Gosner Stage	No Oil - Oil without Plants	-0.01	0.19	5.19	-0.07	1.00
Second Event (GS 40 – 42)	Gosner Stage	Oil with Plants - Oil without Plants	-0.18	0.20	6.66	-0.88	0.67
Second Event (GS 40 – 42)	Snout Vent Length (mm)	No Oil - Oil with Plants	-0.59	0.82	6.10	-0.71	0.77
Second Event (GS 40 – 42)	Snout Vent Length (mm)	No Oil - Oil without Plants	1.02	0.80	5.58	1.28	0.46
Second Event (GS 40 – 42)	Snout Vent Length (mm)	Oil with Plants - Oil without Plants	1.61	0.83	6.31	1.94	0.21
Second Event (GS 40 – 42)	Total Length (mm)	No Oil - Oil with Plants	0.26	2.16	6.05	0.12	0.99
Second Event (GS 40 – 42)	Total Length (mm)	No Oil - Oil without Plants	2.60	2.12	5.65	1.23	0.48
Second Event (GS 40 – 42)	Total Length (mm)	Oil with Plants - Oil without Plants	2.34	2.18	6.26	1.07	0.56

SI Table 3-20 Pooled sample mass for total polycyclic aromatic compound (tPAC) bioaccumulation analysis of tadpole tissues. To achieve sufficient sample mass after removal of GI tract, for the quantification of tPAC bioaccumulation in tadpole tissues, tadpoles were pooled in groups of four (n=4) from the same treatments and same Gosner stage classes

Pool ID (n-4)	Treatment	GS Class	Total Sample Mass (grams)
pool_01	Oil with Plants	GS 34 - 36	1.05
pool_02	Oil with Plants	GS 40 - 42	1.18
pool_03	No Oil	GS 34 - 36	1.05
pool_04	No Oil	GS 40 - 42	1.42
pool_05	Oil without Plants	GS 34 - 36	0.95
pool_06	Oil without Plants	GS 40 - 42	1.19
pool_07	Oil with Plants	GS 34 - 36	1.06
pool_08	Oil with Plants	GS 40 - 42	1.86
pool_09	Oil without Plants	GS 34 - 36	1.13
pool_10	Oil without Plants	GS 40 - 42	1.40
pool_11	Oil without Plants	GS 34 - 36	0.71
pool_12	Oil without Plants	GS 40 - 42	0.84
pool_13	No Oil	GS 34 - 36	1.04
pool_14	No Oil	GS 40 - 42	1.38
pool_15	No Oil	GS 34 - 36	0.83
pool_16	No Oil	GS 40 - 42	1.47
pool_17	Oil with Plants	GS 34 - 36	1.04
pool_18	Oil with Plants	GS 40 - 42	1.55

SI Table 3-21 ANOVA of total, parent and alkylated polycyclic aromatic compound contents (ng/gram of lipid) amongst tadpole tissues. Data from the treatments sampled within each mesocosm at the IISD-Experimental Lakes Area sampled between July 2, 2021 – July 30, 2021, partitioned by sampling event.

Sampling Event (GS Class)	PAH Group	Term	df	SumSQ	MeanSQ	Statistic	P-Value
First event (GS 34-36)	Alkylated PACs	Treatments	2	0.44	0.22	0.272	0.771
First event (GS 34-36)	Alkylated PACs	Residuals	6	4.845	0.808	NA	NA
First event (GS 34-36)	Total PACs	Treatments	2	0.276	0.138	0.241	0.793
First event (GS 34-36)	Total PACs	Residuals	6	3.434	0.572	NA	NA
First event (GS 34-36)	Parent PACs	Treatments	2	0.023	0.011	0.533	0.612
First event (GS 34-36)	Parent PACs	Residuals	6	0.127	0.021	NA	NA
Second Event (GS 40–42)	Alkylated PACs	Treatments	2	0.081	0.041	0.62	0.569
Second Event (GS 40–42)	Alkylated PACs	Residuals	6	0.394	0.066	NA	NA
Second Event (GS 40–42)	Total PACs	Treatments	2	0.143	0.072	1.575	0.282
Second Event (GS 40–42)	Total PACs	Residuals	6	0.273	0.045	NA	NA
Second Event (GS 40–42)	Parent PACs	Treatments	2	0.061	0.031	1.727	0.256
Second Event (GS 40–42)	Parent PACs	Residuals	6	0.106	0.018	NA	NA

SI Table 3-22 Multiple comparisons test of total, parent and alkylated polycyclic aromatic compound contents (ng/gram of lipid) amongst tadpole tissues. Data from the treatments sampled within each mesocosm at the IISD-Experimental Lakes Area sampled between July 2, 2021 - July 30, 2021, partitioned by sampling event.

Sampling Event (GS Class)	PAH Group	Contrast	Estimate	SE	DF	T.Ratio	P.Value
First event (GS 34-36)	Alkylated PACs	No Oil - Oil with Plants	-0.07	0.73	6	-0.10	1.00
First event (GS 34-36)	Alkylated PACs	No Oil - Oil without Plants	-0.50	0.73	6	-0.68	0.78
First event (GS 34-36)	Alkylated PACs	Oil with Plants - Oil without Plants	-0.43	0.73	6	-0.59	0.83
First event (GS 34-36)	Total PACs	No Oil - Oil with Plants	-0.10	0.62	6	-0.16	0.99
First event (GS 34-36)	Total PACs	No Oil - Oil without Plants	-0.41	0.62	6	-0.67	0.79
First event (GS 34-36)	Total PACs	Oil with Plants - Oil without Plants	-0.31	0.62	6	-0.51	0.87
First event (GS 34-36)	Parent PACs	No Oil - Oil with Plants	0.03	0.12	6	0.24	0.97
First event (GS 34-36)	Parent PACs	No Oil - Oil without Plants	-0.09	0.12	6	-0.75	0.75
First event (GS 34-36)	Parent PACs	Oil with Plants - Oil without Plants	-0.12	0.12	6	-0.99	0.61
Second Event (GS 40–42)	Alkylated PACs	No Oil - Oil with Plants	-0.22	0.21	6	-1.07	0.57
Second Event (GS 40–42)	Alkylated PACs	No Oil - Oil without Plants	-0.06	0.21	6	-0.26	0.96
Second Event (GS 40–42)	Alkylated PACs	Oil with Plants - Oil without Plants	0.17	0.21	6	0.81	0.71
Second Event (GS 40–42)	Total PACs	No Oil - Oil with Plants	-0.17	0.17	6	-0.96	0.62
Second Event (GS 40–42)	Total PACs	No Oil - Oil without Plants	0.14	0.17	6	0.81	0.71
Second Event (GS 40–42)	Total PACs	Oil with Plants - Oil without Plants	0.31	0.17	6	1.77	0.26
Second Event (GS 40–42)	Parent PACs	No Oil - Oil with Plants	-0.06	0.11	6	-0.52	0.87
Second Event (GS 40–42)	Parent PACs	No Oil - Oil without Plants	-0.20	0.11	6	-1.80	0.25
Second Event (GS 40–42)	Parent PACs	Oil with Plants - Oil without Plants	-0.14	0.11	6	-1.29	0.45

SI Table 3-23 Mixed Effect Model- ANOVA test on data from weekly behavioural assays conducted on a subset of n=5 tadpoles between July 2, 2021 - July 30, 2021 in the mesocosms, partitioned by study day. Endpoints include mean nearest-neighbour distance (NND), percent time spent within inner sector (%) and total distance.

Study Day	Endpoint	npar	sumsq	meansq	Statistic
7	Total Distance (cm)	2	59811.131	29905.565	1.292
7	Mean NND (cm)	2	3.812	1.906	0.606
7	Percent Time spent in Inner Sector (%)	2	1728.536	864.268	1.835
15	Total Distance (cm)	2	5030.736	2515.368	0.14
15	Mean NND (cm)	2	0.611	0.305	0.058
15	Percent Time spent in Inner Sector (%)	2	328.692	164.346	0.188
21	Total Distance (cm)	2	7372.947	3686.474	0.663
21	Mean NND (cm)	2	27.353	13.677	2.031
21	Percent Time spent in Inner Sector (%)	2	3996.625	1998.312	2.159
28	Total Distance (cm)	2	47452.054	23726.027	1.35
28	Mean NND (cm)	2	10.082	5.041	1.771
28	Percent Time spent in Inner Sector (%)	2	983.257	491.629	0.809
28 (with predator cue)	Total Distance (cm)	2	23975.867	11987.934	0.456
28 (with predator cue)	Mean NND (cm)	2	1.568	0.784	0.488
28 (with predator cue)	Percent Time spent in Inner Sector (%)	2	1320.619	660.309	1.633

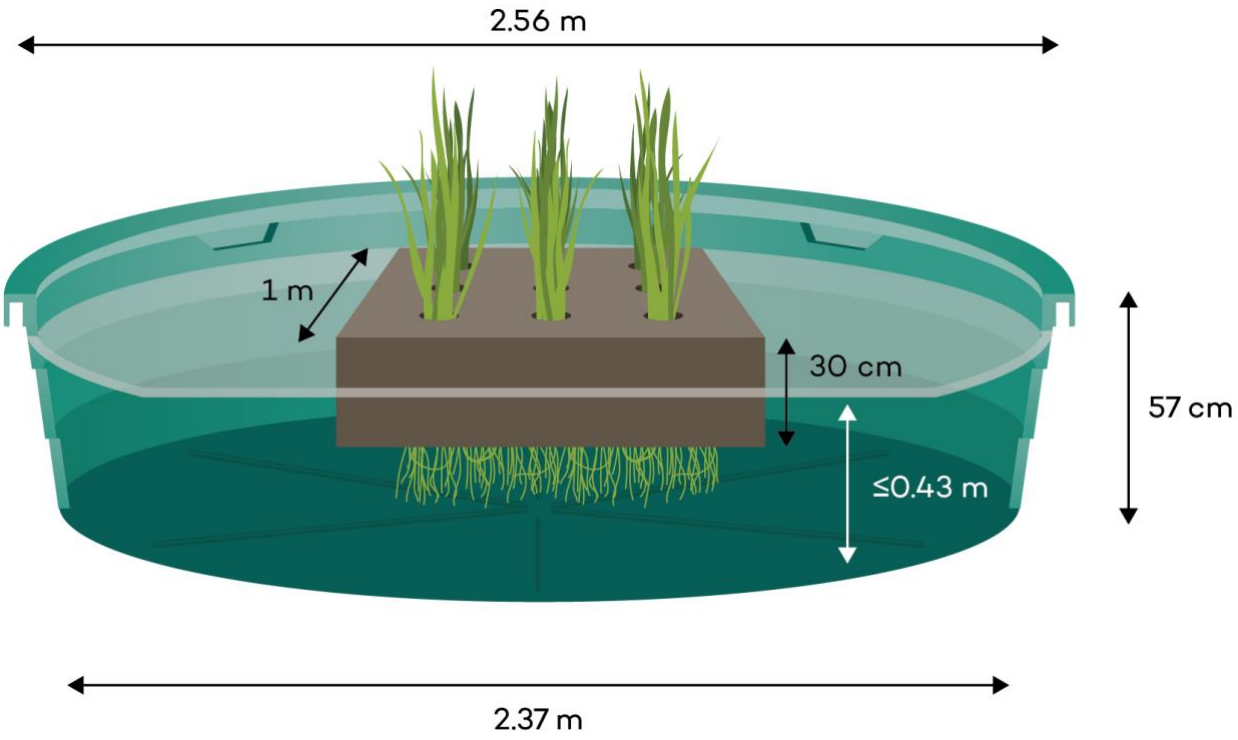
SI Table 3-24 Multiple comparisons test of weekly behavioural assays conducted on a subset of n=5 tadpoles between July 2, 2021 - July 30, 2021 in the mesocosms, partitioned by study day. Statistical analysis of mean nearest neighbour distance (NND), percent time in inner sector (%) and total distance endpoints measured in the behavioural assays revealed no significant differences between treatments.

Study Day	Endpoint	Contrast	Estimate	SE	df	T-Ratio	P-Value
7	Mean NND (cm)	No Oil - Oil with Plants	0.17	0.65	6	0.26	0.97
7	Mean NND (cm)	No Oil - Oil without Plants	-0.52	0.65	6	-0.80	0.72
7	Mean NND (cm)	Oil with Plants - Oil without Plants	-0.68	0.65	6	-1.06	0.57
7	Percent Time Spent in Inner Sector (%)	No Oil - Oil with Plants	-12.39	7.92	6	-1.56	0.33
7	Percent Time Spent in Inner Sector (%)	No Oil - Oil without Plants	1.41	7.92	6	0.18	0.98
7	Percent Time Spent in Inner Sector (%)	Oil with Plants - Oil without Plants	13.80	7.92	6	1.74	0.27
7	Total Distance (cm)	No Oil - Oil with Plants	6.22	64.39	6	0.10	1.00
7	Total Distance (cm)	No Oil - Oil without Plants	92.58	64.39	6	1.44	0.38
7	Total Distance (cm)	Oil with Plants - Oil without Plants	86.36	64.39	6	1.34	0.43
15	Mean NND (cm)	No Oil - Oil with Plants	-0.31	0.90	6	-0.34	0.94
15	Mean NND (cm)	No Oil - Oil without Plants	-0.16	0.90	6	-0.18	0.98
15	Mean NND (cm)	Oil with Plants - Oil without Plants	0.14	0.90	6	0.16	0.99
15	Percent Time Spent in Inner Sector (%)	No Oil - Oil with Plants	4.22	13.11	6	0.32	0.95
15	Percent Time Spent in Inner Sector (%)	No Oil - Oil without Plants	-3.81	13.11	6	-0.29	0.96
15	Percent Time Spent in Inner Sector (%)	Oil with Plants - Oil without Plants	-8.03	13.11	6	-0.61	0.82
15	Total Distance (cm)	No Oil - Oil with Plants	-33.25	104.97	6	-0.32	0.95
15	Total Distance (cm)	No Oil - Oil without Plants	-55.07	104.97	6	-0.53	0.86
15	Total Distance (cm)	Oil with Plants - Oil without Plants	-21.82	104.97	6	-0.21	0.98

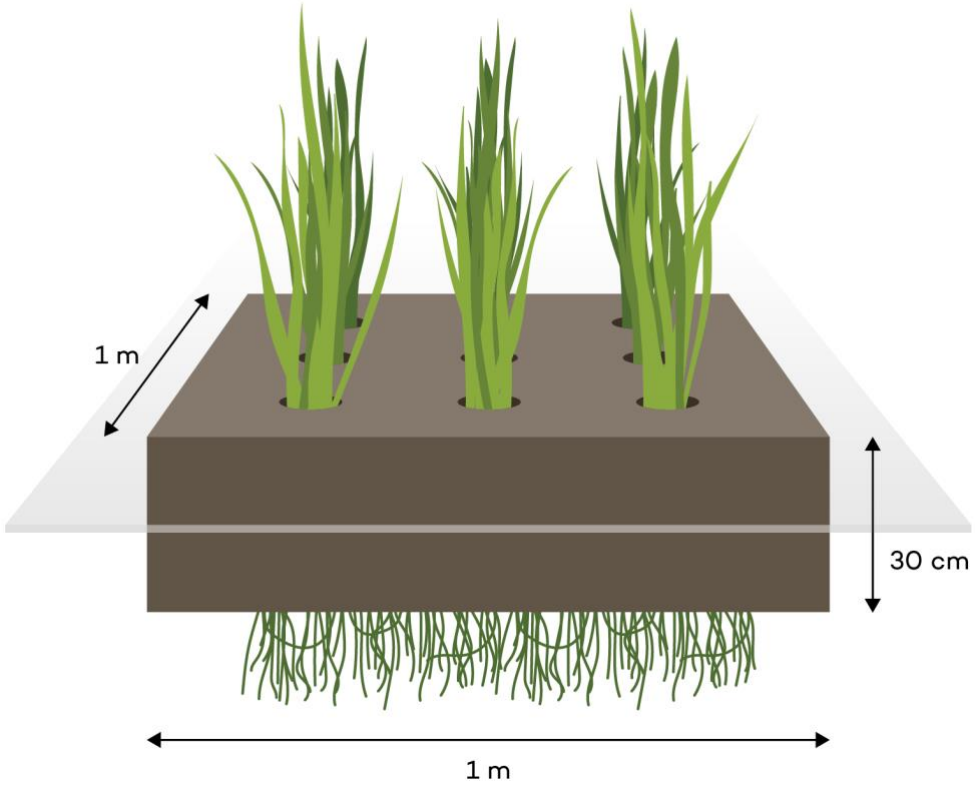
21	Mean NND (cm)	No Oil - Oil with Plants	0.47	0.95	6	0.50	0.88
21	Mean NND (cm)	No Oil - Oil without Plants	1.84	0.95	6	1.94	0.21
21	Mean NND (cm)	Oil with Plants - Oil without Plants	1.37	0.95	6	1.44	0.38
21	Percent Time Spent in Inner Sector (%)	No Oil - Oil with Plants	-2.33	11.11	6	-0.21	0.98
21	Percent Time Spent in Inner Sector (%)	No Oil - Oil without Plants	-21.06	11.11	6	-1.90	0.22
21	Percent Time Spent in Inner Sector (%)	Oil with Plants - Oil without Plants	-18.72	11.11	6	-1.69	0.29
21	Total Distance (cm)	No Oil - Oil with Plants	-46.19	43.74	6	-1.06	0.57
21	Total Distance (cm)	No Oil - Oil without Plants	-40.52	43.74	6	-0.93	0.65
21	Total Distance (cm)	Oil with Plants - Oil without Plants	5.67	43.74	6	0.13	0.99
28	Mean NND (cm)	No Oil - Oil with Plants	-0.88	0.62	6	-1.43	0.39
28	Mean NND (cm)	No Oil - Oil without Plants	-1.09	0.62	6	-1.78	0.26
28	Mean NND (cm)	Oil with Plants - Oil without Plants	-0.21	0.62	6	-0.35	0.94
28	Percent Time Spent in Inner Sector (%)	No Oil - Oil with Plants	3.08	9.87	6	0.31	0.95
28	Percent Time Spent in Inner Sector (%)	No Oil - Oil without Plants	-9.00	9.87	6	-0.91	0.65
28	Percent Time Spent in Inner Sector (%)	Oil with Plants - Oil without Plants	-12.08	9.87	6	-1.22	0.48
28	Total Distance (cm)	No Oil - Oil with Plants	23.79	74.02	6	0.32	0.95
28	Total Distance (cm)	No Oil - Oil without Plants	-91.40	74.02	6	-1.24	0.48
28	Total Distance (cm)	Oil with Plants - Oil without Plants	-115.19	74.02	6	-1.56	0.33
28 (predator cue)	Mean NND (cm)	No Oil - Oil with Plants- Predator Cue	0.46	0.46	6	0.98	0.61
28 (predator cue)	Mean NND (cm)	No Oil - Oil without Plants - Predator Cue	0.27	0.46	6	0.58	0.84
28 (predator cue)	Mean NND (cm)	Oil with Plants - Oil without Plants - Predator Cue	-0.19	0.46	6	-0.40	0.92
28 (predator cue)	Percent Time Spent in Inner Sector (%)	No Oil - Oil with Plants- Predator Cue	-5.07	7.34	6	-0.69	0.78

28 (predator cue)	Percent Time Spent in Inner Sector (%)	No Oil - Oil without Plants - Predator Cue	-13.16	7.34	6	-1.79	0.25
28 (predator cue)	Percent Time Spent in Inner Sector (%)	Oil with Plants - Oil without Plants - Predator Cue	-8.09	7.34	6	-1.10	0.55
28 (predator cue)	Total Distance (cm)	No Oil - Oil with Plants- Predator Cue	-4.23	117.03	6	-0.04	1.00
28 (predator cue)	Total Distance (cm)	No Oil - Oil without Plants - Predator Cue	94.58	117.03	6	0.81	0.71
28 (predator cue)	Total Distance (cm)	Oil with Plants - Oil without Plants - Predator Cue	98.81	117.03	6	0.84	0.69

SI Figure 3-1 Aerial image of mesocosm site located at IISD – Experimental Lakes Area (49°42'20.3"N 93°45'50.9"W). Mean water volume for the mesocosms (n=9) was 1,531.8 L (SD=138.1).

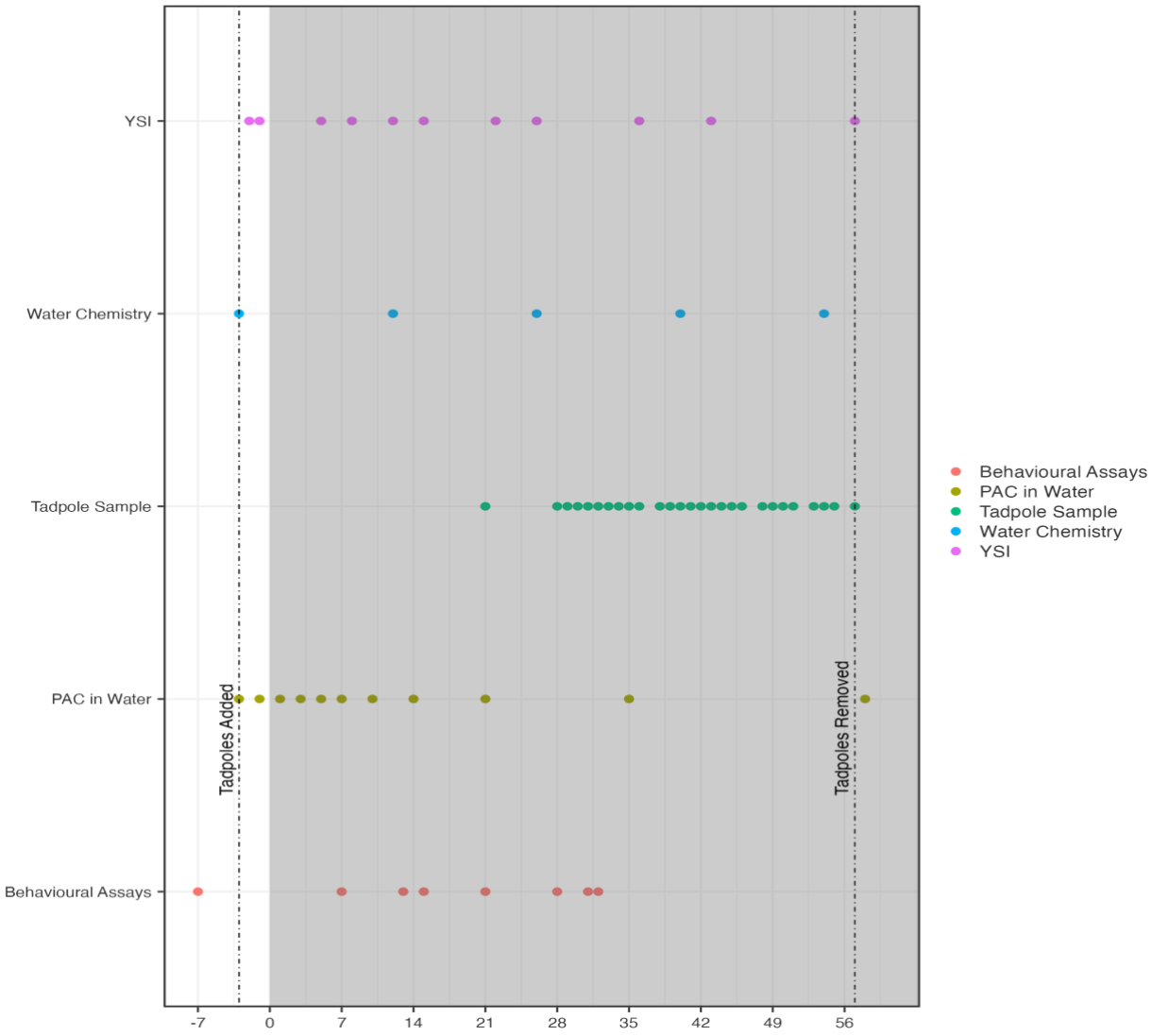


SI Figure 3-2 Schematic of engineered floating wetland present (planted) in the mesocosms located at IISD – Experimental Lakes Area. Mean water volume for the mesocosms (n=9) was 1,531.8 L (SD=138.1).

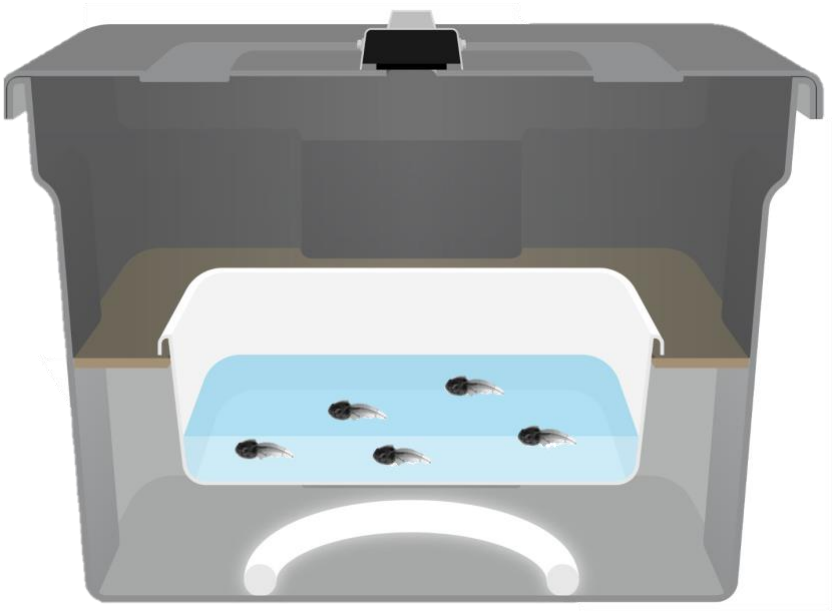


SI Figure 3-3 Schematic of engineered floating wetlands present (either planted or unplanted) applied in the mesocosms.

SI Figure 3-4 A large 600 litre plastic cattle tank filled ¾ with water, and then within each larger tank, three 70 litre tanks, filled ¾ with Lake 260 water. Each smaller 70 Litre tank covered by predator mesh with aeration tanking submerged in the water. HOBO temperature loggers are present in each of the 70 litre tanks and these tanks contained the wood frog egg masses.



SI Figure 3-5 Sampling and monitoring schedule. Grayed area represents CHV-WAF addition within treatment mesocosms.



SI Figure 3-6 Behavioural Arena schematic. Behavioural assays conducted weekly on subset of 5 tadpoles.

Chapter 4 - Synthesis and Conclusions

The findings of this thesis provide a comprehensive examination of the effects of exposure to conventional heavy crude oil (CHV) – derived hydrocarbons on wood frog (*Lithobates sylvaticus*) tadpoles across a range of endpoints, including developmental, histological, bioaccumulation, and behavioural responses. By using in-lake (enclosures) and land-based (tank) mesocosms (Chapter 2 and Chapter 3 respectively), these two studies characterized potential impacts of CHV contamination on amphibian populations and contribute valuable insights into ecological risk assessment in boreal freshwater systems. In this chapter we synthesize our findings, make recommendations, acknowledge limitations and propose future work.

The main objective of Chapter 2 was **to characterize the effects of chronic exposure to conventional heavy crude oil–derived hydrocarbons on the growth, development, histopathology, and bioaccumulation in wood frogs in a freshwater ecosystem**. Tadpoles exposed to CHV within shoreline enclosures did not show significant adverse effects across developmental and bioaccumulation endpoints at the exposure levels tested. While tPACs accumulated in tadpole tissue initially, these levels declined over time, suggesting an ability in tadpoles to metabolize or eliminate certain polycyclic aromatic compounds (PACs) under moderate CHV concentrations. It was hypothesized that exposure could have an impact on developmental parameters and bioaccumulation potentially affecting long-term survivability if exposed early in wood frog tadpole early larval stages. The null hypothesis posited that there would be no significant differences in developmental parameters and bioaccumulation of CHV-derived hydrocarbons in wood frogs resulting from chronic exposure during larval stages. The lack of persistent negative effects on the previously mentioned endpoints as well as liver histology does not support this hypothesis and suggests that wood frogs may tolerate moderate CHV contamination during early developmental stages. Importantly, this chapter underscores the potential resilience of wood frogs to early life-stage CHV exposure, which may extend to other amphibian species in similar environments. However, long-term ecological effects, especially on higher trophic levels, remain unknown.

Chapter 3 extended the analysis to in-land mesocosms and added behavioural assays to assess whether CHV-derived hydrocarbons (via a WAF) influenced tadpole activity. **The objective of the study was to characterize effects of chronic exposure to CHV-derived hydrocarbons on the growth and development, behaviour, histopathology, and**

bioaccumulation in wood frogs in a freshwater ecosystem. No measurable stress or behavioural changes were observed in tadpoles exposed to minimal CHV – derived hydrocarbon concentrations, suggesting effective PAC metabolization and limited exposure over time. While no significant differences were detected in liver histology or other growth parameters, behavioural assays did confirm expected responses in ethanol (anxiolytic) and caffeine (anxiogenic) tests, underscoring the sensitivity of the test system and analysis methodologies and their potential utility for future toxicity studies in field settings. The results suggest that tadpoles did not experience measurable stress or altered behaviour due to the CHV-WAF exposure, under the study conditions. The findings also suggest that, at minimal CHV-WAF concentrations, wood frog tadpoles did not experience critical developmental or behavioural disruptions, due to the efficient metabolization of PACs and low, declining CHV-WAF exposure levels over time.

Overall, there was no apparent difference between reference and oiled treatments in the histological appearance of the liver in Chapter 2 and in Chapter 3, there was also no apparent difference among treatment groups in the histological appearance of the liver. Irrespective of experimental treatments, there are apparent differences in liver histology between the two studies. For example, the cords of hepatocytes in the Chapter 2 *in-situ* enclosures were interspersed with sinusoids (less prominent than in the mesocosms of Chapter 3) and blood vessels containing erythrocytes. Control samples from the *in-situ* enclosures appear to have much larger hepatocytes with more abundant cytoplasmic vacuolation than those from the mesocosm study. Hepatocyte vacuolation typically represents stored energy products (i.e. likely glycogen and/or lipids) and may reflect better overall condition of tadpoles in the *in-situ* enclosure study. Consistent results from Regnault *et al.* (2016); Regnault *et al.* (2018) showed an increase in lipid droplet content within hepatocytes, along with structural alterations in the parenchyma and irregular cell shapes (Regnault *et al.*, 2016; 2018). In both Chapter 2 and Chapter 3 studies, hepatocyte cords were interspersed with sinusoids and blood vessels, which contained erythrocytes, indicative of normal histological structure across both treatments. This can be compared to increased lipid droplets in liver cells, and decreased pigment content when conducting histological structures of the liver, indicative of hepatic stress and have been documented in western clawed frog tadpoles exposed to Benzo(a)pyrene (Regnault et al., 2014). Note that mean tadpole weight in reference groups was markedly different in the two studies; tanks with tadpoles that were free feeding had a mean weight of 0.80 grams compared to the tadpoles within the enclosures that were fed every second day had

a mean weight of 1.95 grams. Water quality and environmental factors, including temperature differences between tanks (21.5°C) and *in-situ* enclosures (23.3°C), also could have influenced development, highlighting the challenges of direct comparisons between these two studies. Additionally, one study contained tadpoles within tanks, meanwhile one study contained tadpoles within cages within a large in lake shoreline enclosure, furthering the idea that comparisons between the studies for this element was not a simple comparison.

The combined results from Chapters 2 and 3 provide a comprehensive view of the effects of chronic exposure to conventional heavy hydrocarbons derived from conventional crude oil (CHV) derived hydrocarbons on wood frog tadpoles across multiple levels of biological organization in boreal freshwater systems. Chapter 2 explored developmental, histological, and bioaccumulation responses, revealing that at the CHV concentrations tested within the shoreline enclosures, wood frog tadpoles did not exhibit significant adverse effects. Although there was an initial increase in tPACs within tadpole tissues, these levels declined over the duration of the exposure, suggesting an ability in tadpoles to metabolize or eliminate certain PACs. Histological examinations further supported this, as there were no apparent liver abnormalities, and developmental endpoints like growth were largely unaffected. This suggests that tadpoles can withstand moderate levels of CHV derived hydrocarbon exposure during critical early life stages without marked physiological consequences.

Chapter 3 extended the analysis to mesocosms, introducing behavioural assays to assess whether CHV-derived hydrocarbons from a water-accommodated fraction could influence tadpole activity. The results showed that tadpoles in CHV derived hydrocarbons (WAF) conditions did not experience measurable stress, altered behaviour, or disruptions in growth metrics, even in the presence of predator cues. While behavioural assays detected expected responses to ethanol (anxiolytic) and caffeine (anxiogenic) controls, indicating the sensitivity of the methods, there was no significant effect from CHV-WAF exposure on tadpole behaviour. Together, these findings suggest that low-level CHV derived hydrocarbon exposure in natural environments does not critically impact wood frog tadpoles across developmental, survival, physiological, and behavioural metrics. Notably, these findings contrast with existing literature on the toxicity of polycyclic aromatic compounds (PACs) in freshwater amphibians, likely due to the low exposure levels in this study. Overall, these results given the low levels of PAC exposure throughout the

tank study, underscore the resilience of wood frog tadpoles to CHV exposure at early developmental stages and highlight the importance of examining biological responses at multiple levels to accurately assess and characterize the ecological impact of CHV contamination on amphibians.

4.1 Limitation and Challenges

One limitation of the *in-situ* enclosure study (Chapter 2) was the restricted sampling frequency, which reduced comparability between the two studies, particularly for endpoints which required repeated measures of the tadpoles, such as growth rate. Lethal sampling was conducted at specific Gosner stages to minimize handling stress on tadpoles. However, more frequent non-lethal sampling could improve growth rate analysis.

Another limitation for the Chapter 2 experiment was the lack of behavioural assays conducted. Logistical constraints prevented robust behavioural assays for the *in-lake* enclosures. Weekly behavioural assays would enhance the characterization of CHV exposure, however, based on the field study location and with the decision to prevent undue stress on the research subjects, it was determined behavioural assays were not achievable for the experiment described in Chapter 2.

The tank study (Chapter 3) also had limitations that may have influenced the observed results, including low initial PAC concentrations, sorption to mesocosm materials, and dilution from water replenishment. These factors effectively reduced WAF levels to non-detectable levels. As a result, exposure levels in the mesocosms were minimal after just a few days. Given that tadpoles are known to efficiently metabolize certain PACs, the concentrations during the early life stages were likely too low to elicit observable toxicological effects.

In terms of the experimental design in the tank experiment, a key limitation was the absence of a treatment with no oil and no plants. As discussed in Chapter 3, tadpoles in the oil only treatment (without plants) showed slightly poorer condition, which was likely due to limited food resources compared to treatments with plants. Plants in the EFWs provided greater amounts of biofilm from their roots, serving as an additional food source. Unfortunately, without a no oil & no plants treatment, it was impossible to distinguish the effects of limited food availability from

those of oil exposure or determine whether resource limitation exacerbated the tadpoles' sensitivity to oil. Resource constraints and the pre-existing presence of EFWs (with vegetation planted a year prior during the global COVID-19 pandemic) prevented the inclusion of this treatment.

Another limitation of the study was the construction material of the tanks. Evidence from other aspects of the larger FloWTER project suggested that high-density polyethylene (HDPE) mesocosms may have adsorbed PACs from the WAF, further reducing exposure levels and toxicological effects. Consequently, PAC concentrations in the treatment mesocosms approached baseline levels after approximately 10 days of exposure.

To better characterize the effects of CHV exposure, future studies should replicate similar study parameters but test significantly higher CHV and WAF concentrations. It is important to note that the WAF in this study was prepared by exposing a relatively large amount of oil to a small volume of water for an extended duration (~36 hours compared to the 24 hours typical in lab-prepared WAFs). This approach was chosen as it reflects more realistic environmental conditions than the highly mixed, "high-energy" WAFs commonly created in laboratory settings. Furthermore, the low PAC levels observed in the WAF in Chapter 3 were likely due to minimal mixing during preparation, leading to what is referred to as a "low-energy WAF." While this method introduces certain limitations, it provides a more accurate simulation of natural conditions compared to the standardized 24-hour mixing protocols frequently employed in controlled lab studies.

4.2 Overall Conclusions and Future Research

Future research should address these limitations and investigate higher CHV exposure levels, while remaining within realistic scenarios. This study highlights the importance of understanding amphibian responses to freshwater oil contamination, which is essential for ecological risk assessments of potential future spills. Although wood frog tadpoles displayed resilience to moderate CHV exposure, further research is needed to examine long-term and multi-generational impacts, particularly at higher contamination levels while the limitations within the study are addressed. The observed bioaccumulation, even if transient, highlights the need to consider

potential long-term ecological implications, especially as bioaccumulated tPACs in tadpoles could affect higher trophic levels. Additionally, the absence of higher WAF concentrations (tPAC exposure) that were minimized over time may limit the scope of the conclusions, as potential toxicological effects may only manifest under more substantial contamination.

REFERENCES

Afzal, M., Rehman, K., Shabir, G., Tahseen, R., Ijaz, A., Hashmat, A.J., Brix, H., 2019. Large-scale remediation of oil-contaminated water using floating treatment wetlands. npj Clean Water 2. 10.1038/s41545-018-0025-7

Ågerstrand, M., Arnold, K., Balshine, S., Brodin, T., Brooks, B.W., Maack, G., McCallum, E.S., Pyle, G., Saaristo, M., Ford, A.T., 2020. Emerging investigator series: use of behavioural endpoints in the regulation of chemicals. Environmental science--processes & impacts 22, 49-65. 10.1039/c9em00463g

Alcaraz, A.J.G., Baraniuk, S., Mikulášek, K., Park, B., Lane, T., Burbridge, C., Ewald, J., Potěšil, D., Xia, J., Zdráhal, Z., Schneider, D., Crump, D., Basu, N., Hogan, N., Brinkmann, M., Hecker, M., 2022. Comparative analysis of transcriptomic points-of-departure (tPODs) and apical responses in embryo-larval fathead minnows exposed to fluoxetine. Environmental Pollution 295, 118667. https://doi.org/10.1016/j.envpol.2021.118667

Alcaraz, A.J.G., Mikulášek, K., Potěšil, D., Park, B., Shekh, K., Ewald, J., Burbridge, C., Zdráhal, Z., Schneider, D., Xia, J., Crump, D., Basu, N., Hecker, M., 2021a. Assessing the Toxicity of 17α-Ethinylestradiol in Rainbow Trout Using a 4-Day Transcriptomics Benchmark Dose (BMD) Embryo Assay. Environmental Science & Technology 55, 10608-10618. 10.1021/acs.est.1c02401

Alcaraz, A.J.G., Potěšil, D., Mikulášek, K., Green, D., Park, B., Burbridge, C., Bluhm, K., Soufan, O., Lane, T., Pipal, M., Brinkmann, M., Xia, J., Zdráhal, Z., Schneider, D., Crump, D., Basu, N., Hogan, N., Hecker, M., 2021b. Development of a Comprehensive Toxicity Pathway Model for 17α-Ethinylestradiol in Early Life Stage Fathead Minnows (Pimephales promelas). Environmental Science & Technology 55, 5024-5036. 10.1021/acs.est.0c05942

Alford, R.A., Richards, S.J., 1999. Global amphibian declines: A problem in applied ecology. Annu. Rev. Ecol. Syst. 30, 133-165. 10.1146/annurev.ecolsys.30.1.133

Allaire, J.J., Xie, Y., Dervieux, C., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., Chang, W., Iannone, R., 2023. rmarkdown: Dynamic Documents for R. https://github.com/rstudio/rmarkdown

Alsaadi, F., Hodson, P.V., Langlois, V.S., 2018. An Embryonic Field of Study: The Aquatic Fate and Toxicity of Diluted Bitumen. Bulletin of Environmental Contamination and Toxicology 100, 8-13. 10.1007/s00128-017-2239-7

Audo, M.C., Mann, T.M., Polk, T.L., Loudenslager, C.M., Diehl, W.J., Altig, R., 1995. Food Deprivation during Different Periods of Tadpole (Hyla chrysoscelis) Ontogeny Affects Metamorphic Performance Differently. Oecologia 103, 518-522. 10.1007/BF00328691

Awkerman, J.A., Raimondo, S., 2018. Simulated developmental and reproductive impacts on amphibian populations and implications for assessing long-term effects. Ecotoxicology and environmental safety 149, 233-240. 10.1016/j.ecoenv.2017.11.047

Ball, A., Truskewycz, A., 2013. Polyaromatic hydrocarbon exposure: an ecological impact ambiguity. Environmental Science and Pollution Research 20, 4311-4326. 10.1007/s11356-013-1620-2

Balmer, J.E., Hung, H., Yu, Y., Letcher, R.J., Muir, D.C.G., 2019. Sources and environmental fate of pyrogenic polycyclic aromatic hydrocarbons (PAHs) in the Arctic. Emerging Contaminants 5, 128-142. 10.1016/j.emcon.2019.04.002

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67, 1-48. 10.18637/jss.v067.i01

Bates, D., Maechler, M., Bolker, B., Walker, S., 2023. lme4: Linear Mixed-Effects Models using Eigen and S4. https://github.com/lme4/lme4/

Belz, R.G., Piepho, H.P., Wu, R., 2012. Modeling effective dosages in hormetic dose-response studies. PloS one 7, e33432. 10.1371/journal.pone.0033432

Bertram, M.G., Ågerstrand, M., Thoré, E.S.J., Allen, J., Balshine, S., Brand, J.A., Brooks, B.W., Dang, Z., Duquesne, S., Ford, A.T., Hoffmann, F., Hollert, H., Jacob, S., Kloas, W., Klüver, N., Lazorchak, J., Ledesma, M., Maack, G., Macartney, E.L., Martin, J.M., Melvin, S.D., Michelangeli, M., Mohr, S., Padilla, S., Pyle, G., Saaristo, M., Sahm, R., Smit, E., Steevens, J.A., van den Berg, S., Vossen, L.E., Wlodkowic, D., Wong, B.B.M., Ziegler, M., Brodin, T., 2024. EthoCRED: a framework to guide reporting and evaluation of the relevance and reliability of behavioural ecotoxicity studies. Biological Reviews n/a. https://doi.org/10.1111/brv.13154

Bilodeau, J.C., Villagomez, J.M.G., Kimpe, L.E., Thomas, P.J., Pauli, B.D., Trudeau, V.L., Blais, J.M., 2019. Toxicokinetics and bioaccumulation of polycyclic aromatic compounds in wood frog tadpoles (Lithobates sylvaticus) exposed to Athabasca oil sands sediment. Aquatic Toxicology 207, 217-225. 10.1016/j.aquatox.2018.11.006

Blaustein, A.R., Romansic, J.M., Kiesecker, J.M., Hatch, A.C., 2003. Ultraviolet radiation, toxic chemicals and amphibian population declines. Diversity and Distributions 9, 123-140. 10.1046/j.1472-4642.2003.00015.x

Bouffard, J., 2021. Effects of a Neonicotinoid Insecticide and Population Density on Behaviour and Development of Wood Frogs (Lithobates sylvaticus). Université d'Ottawa/University of Ottawa.

Bradley, E.C., Tracy, L., 2013. Personality Traits Are Expressed in Bullfrog Tadpoles during Open-Field Trials. Journal of herpetology 47, 378-383. 10.1670/12-061

Bridges, C.M., 1997. Tadpole swimming performance and activity affected by acute exposure to sublethal levels of carbaryl. Environmental Toxicology and Chemistry 16, 1935-1939. 10.1002/etc.5620160924

Bryer, P.J., Elliott, J.N., Willingham, E.J., 2006. The Effects of Coal Tar Based Pavement Sealer on Amphibian Development and Metamorphosis. Ecotoxicology 15, 241-247. 10.1007/s10646-005-0055-z

Bulaeva, E., Lanctôt, C., Reynolds, L., Trudeau, V.L., Navarro-Martín, L., 2015. Sodium perchlorate disrupts development and affects metamorphosis- and growth-related gene expression in tadpoles of the wood frog (Lithobates sylvaticus). General and comparative endocrinology 222, 33-43. 10.1016/j.ygcen.2015.01.012

Canada, N.-R.-. 2020. Energy Fact Book 2020-2021. https://www.nrcan.gc.ca/science-data/data-analysis/energy-data-analysis/energy-facts/20061

Canada, R.A.o., 2019. Canadian Crude Oil Transportation: Comparing the Safety of Pipelines and Railways. https://www.railcan.ca/wp-content/uploads/2019/08/Canadian-Crude-Oil-Transportation-Comparing-the-Safety-of-Pipelines-and-Railways-2.pdf

Careau, V., Garland, T., 2012. Performance, Personality, and Energetics: Correlation, Causation, and Mechanism. Physiological and Biochemical Zoology 85, 543-571. 10.1086/666970

Carlson, B.E., Langkilde, T., 2013. Personality Traits Are Expressed in Bullfrog Tadpoles during Open-Field Trials. Journal of Herpetology 47, 378-383. http://www.jstor.org.uml.idm.oclc.org/stable/43287357

Cedergreen, N., Ritz, C., Streibig, J.C., 2005. Improved empirical models describing hormesis. Environmental toxicology and chemistry 24, 3166-3172. 10.1897/05-014R.1

Colville, C., Alcaraz, A.J., Green, D., Park, B., Xia, J., Soufan, O., Hruška, P., Potěšil, D., Zdráhal, Z., Crump, D., Basu, N., Hogan, N., Hecker, M., 2022. Characterizing toxicity pathways of fluoxetine to predict adverse outcomes in adult fathead minnows (Pimephales promelas). Science of The Total Environment 817, 152747. https://doi.org/10.1016/j.scitotenv.2021.152747

Couillard, C.M., Lee, K., Légaré, B., King, T.L., 2005. Effect of dispersant on the composition of the water-accommodated fraction of crude oil and its toxicity to larval marine fish. Environmental Toxicology and Chemistry 24, 1496. 10.1897/04-267r.1

Currie, Z., 2018. Assessing The Toxicity of a Petroleum-based Hydraulic Oil to Aquatic Organisms and the Photo-induced Toxicity of Polycyclic Aromatic Hydrocarbons in Two Amphibian Species.

Denoël, M., Bichot, M., Ficetola, G.F., Delcourt, J., Ylieff, M., Kestemont, P., Poncin, P., 2010. Cumulative effects of road de-icing salt on amphibian behavior. Aquatic toxicology 99, 275-280. 10.1016/j.aquatox.2010.05.007

Denoël, M., D'Hooghe, B., Ficetola, G.F., Brasseur, C., De Pauw, E., Thomé, J.-P., Kestemont, P., 2012. Using sets of behavioral biomarkers to assess short-term effects of pesticide: a study case with endosulfan on frog tadpoles. Ecotoxicology 21, 1240-1250. 10.1007/s10646-012-0878-3

Denton, D., 2019. Behavioural Responses of Wood Frog (Lithobates Sylvaticus) Tadpoles to Diluted Bitumen. University of Winnipeg.

Dodd, C.K., Jr., 2013. Frogs of the United States and Canada. Johns Hopkins University Press, Baltimore.

Egan, R.J., Bergner, C.L., Hart, P.C., Cachat, J.M., Canavello, P.R., Elegante, M.F., Elkhayat, S.I., Bartels, B.K., Tien, A.K., Tien, D.H., Mohnot, S., Beeson, E., Glasgow, E., Amri, H., Zukowska, Z., Kalueff, A.V., 2009. Understanding behavioral and physiological phenotypes of stress and anxiety in zebrafish. Behavioural Brain Research 205, 38-44. 10.1016/j.bbr.2009.06.022

Egea-Serrano, A., Relyea, R.A., Tejedo, M., Torralva, M., 2012. Understanding of the impact of chemicals on amphibians: a meta-analytic review. Ecology and evolution 2, 1382-1397. 10.1002/ece3.249

Fathallah, S., Medhioub, M.N., Kraiem, M.M., 2012. Photo-induced Toxicity of Four Polycyclic Aromatic Hydrocarbons (PAHs) to Embryos and Larvae of the Carpet Shell Clam Ruditapes decussatus. Bulletin of Environmental Contamination and Toxicology 88, 1001-1008. 10.1007/s00128-012-0603-1

Ferrari, M.C.O., Messier, F., Chivers, D.P., 2008. Larval amphibians learn to match antipredator response intensity to temporal patterns of risk. Behavioral Ecology 19, 980-983. 10.1093/beheco/arn056

Firke, S., 2023. janitor: Simple Tools for Examining and Cleaning Dirty Data. https://CRAN.R-project.org/package=janitor

Garnier, S., 2023. viridis: Colorblind-Friendly Color Maps for R. https://CRAN.R-project.org/package=viridis

Genz, A., Bretz, F., Miwa, T., Mi, X., Hothorn, T., 2023. mvtnorm: Multivariate Normal and t Distributions. http://mvtnorm.R-forge.R-project.org

Gosner, K.L., 1960. A Simplified Table for Staging Anuran Embryos and Larvae with Notes on Identification. Herpetologica 16, 183-190. www.jstor.org/stable/3890061

Grant, E.H.C., 2008. Visual Implant Elastomer Mark Retention Through Metamorphosis in Amphibian Larvae. Journal of Wildlife Management 72, 1247-1252. 10.2193/2007-183

Graves, S., Piepho, H.-P., from Sundar Dorai-Raj, L.S.W.H., 2024. multcompView: Visualizations of Paired Comparisons. https://CRAN.R-project.org/package=multcompView

Green, D.M., Weir, L.A., Casper, G.S., Lannoo, M.J., 2013. Family Ranidae. North American Amphibians. University of California Press, pp. 94-123. http://www.jstor.org/stable/10.1525/j.ctt5hjj7h.14

Havens, S.M., Cooney, B.T., Stainton, M.P., 2024. The chemical analysis of fresh water. International Institute for Sustainable Development 3rd.

Hedtke, S.F., Puglisi, F.A., 1982. Short-term toxicity of five oils to four freshwater species. Archives of Environmental Contamination and Toxicology 11, 425-430. 10.1007/BF01056068

Henry, J., Rodriguez, A., Wlodkowic, D., 2019. Impact of digital video analytics on accuracy of chemobehavioural phenotyping in aquatic toxicology. PeerJ 7, e7367. 10.7717/peerj.7367

Herreid Ii, C.F., Kinney, S., 1967. Temperature and Development of the Wood Frog, Rana Sylvatica, in Alaska. Ecology 48, 579-590. 10.2307/1936502

Hersikorn, B.D., Ciborowski, J.J.C., Smits, J.E.G., 2010. The effects of oil sands wetlands on wood frogs (Rana sylvatica). Toxicological & Environmental Chemistry 92, 1513-1527. 10.1080/02772240903471245

Hersikorn, B.D., Smits, J.E.G., 2011. Compromised metamorphosis and thyroid hormone changes in wood frogs (Lithobates sylvaticus) raised on reclaimed wetlands on the Athabasca oil sands. Environmental Pollution 159, 596-601. 10.1016/j.envpol.2010.10.005

Hopkins, W.A., 2007. Amphibians as Models for Studying Environmental Change. ILAR Journal 48, 270-277. 10.1093/ilar.48.3.270

Hothorn, T., Bretz, F., Westfall, P., 2008. Simultaneous Inference in General Parametric Models. Biometrical Journal 50, 346-363.

Hussar, E., Richards, S., Lin, Z.-Q., Dixon, R.P., Johnson, K.A., 2012. Human Health Risk Assessment of 16 Priority Polycyclic Aromatic Hydrocarbons in Soils of Chattanooga, Tennessee, USA. Water, Air, & Soil Pollution 223, 5535-5548. 10.1007/s11270-012-1265-7

Iannella, M., Liberatore, L., Biondi, M., 2017. Marking tadpoles with Visible Implant Elastomer (VIE) tags: methods for improving readability and decreasing mortality. Salamandra 53, 531-536.

Idowu, I., Francisco, O., Thomas, P.J., Johnson, W., Marvin, C., Stetefeld, J., Tomy, G.T., 2018. Validation of a simultaneous method for determining polycyclic aromatic compounds and alkylated isomers in biota. Rapid Communications in Mass Spectrometry 32, 277-287. 10.1002/rcm.8035

Jofre, M.B., Rosenshield, M.L., Karasov, W.H., 2000. Effects of PCB 126 and Ammonia, Alone and in Combination, on Green Frog (Rana clamitans) and Leopard Frog (R. pipiens) Hatching Success, Development, and Metamorphosis. The Journal of the Iowa Academy of Science 107, 113-122.

Johansson, A., Nilsson, A.N., 1992. Dytiscus latissimus and D. circumcinctus (Coleoptera, Dytiscidae) larvae as predators on three case-making caddis larvae. Hydrobiologia 248, 201-213. 10.1007/bf00006148

José Pinheiro, Douglas Bates, Team., R.C., 2024. nlme: Linear and Nonlinear Mixed Effects Models. https://CRAN.R-project.org/package=nlme

Kharey, G.S., Palace, V., Whyte, L., Greer, C.W., 2024. Influence of heavy Canadian crude oil on pristine freshwater boreal lake ecosystems in an experimental oil spill. FEMS Microbiology Ecology 100. 10.1093/femsec/fiae054

Kheraj, S., 2020. A History of Oil Spills on Long-Distance Pipelines in Canada. The Canadian historical review 101, 161-191. 10.3138/chr.2019-0005

Klemish, J.L., Bogart, S.J., Luek, A., Lannoo, M.J., Pyle, G.G., 2018. Nickel toxicity in wood frog tadpoles: Bioaccumulation and sublethal effects on body condition, food consumption, activity, and chemosensory function. Environmental Toxicology and Chemistry 37, 2458-2466. 10.1002/etc.4210

Krohn, R.M., Palace, V., Smits, J.E.G., 2020. Metal Changes in Pre- and Post-metamorphic Wood Frog (Lithobates sylvaticus) Tadpoles: Implications for Ecotoxicological Studies. Archives of environmental contamination and toxicology. 10.1007/s00244-020-00735-w

Lara-Jacobo, L.R., Willard, B., Wallace, S.J., Langlois, V.S., 2019. Cytochrome P450 1A transcript is a suitable biomarker of both exposure and response to diluted bitumen in developing frog embryos. Environmental Pollution 246, 501-508. 10.1016/j.envpol.2018.12.039

Lee, K., Boufadel, M., Chen, B., Foght, J., Hodson, P., Swanson, S., Venosa, A., 2015. The Royal Society of Canada Expert Panel : The Behaviour and Environmental Impacts of Crude Oil Released into Aqueous Environments. Royal Society of Canada, Ottawa, ON, CA.

Lee, K., Boufadel, M., Chen, B., Foght, J., Hodson, P., Swanson, S., Venosa, A., 2016. High-Priority Research Needs for Oil Spills in Canada: Summary of a Royal Society Expert Panel Report on the Behaviour and Environmental Impacts of Crude Oil Released into Aqueous Environments.

Lee-Jenkins, S.S.Y., Robinson, S.A., 2018. Effects of neonicotinoids on putative escape behavior of juvenile wood frogs (Lithobates sylvaticus) chronically exposed as tadpoles. Environmental Toxicology and Chemistry 37, 3115-3123. 10.1002/etc.4284

Lenth, R.V., 2024. emmeans: Estimated Marginal Means, aka Least-Squares Means. https://CRAN.R-project.org/package=emmeans

Marappa Gounder, R., 2019. Introductory Chapter: Heavy Crude Oil Processing - An Overview. IntechOpen. 10.5772/intechopen.90425

Marquis, O., Millery, A., Guittonneau, S., Miaud, C., 2006. Solvent toxicity to amphibian embryos and larvae. Chemosphere 63, 889-892. https://doi.org/10.1016/j.chemosphere.2005.07.063

Mirza, R.S., Laraby, C.A., Marcellus, A.M., 2013. Knowing Your Behaviour: The importance of Behavioural Assays in the Characterisation of Chemical Alarm Cues in Fishes and Amphibians. Springer New York, pp. 295-308. 10.1007/978-1-4614-5927-9_24

Moeun, B., 2018. Toxicokinetics and Bioaccumulation of Metals in Wood Frog Tadpoles (Lithobates sylvaticus) Exposed to Sediment Near Oil Sands Mining in Northern Alberta. 10.20381/ruor-22411

Mouche, I., Malesic, L., Gillardeaux, O., 2011. FETAX Assay for Evaluation of Developmental Toxicity. Humana Press, pp. 257-269. 10.1007/978-1-60761-849-2_15

Müller, K., 2020. here: A Simpler Way to Find Your Files. https://CRAN.R-project.org/package=here

Mundy, L.J., Bilodeau, J.C., Schock, D.M., Thomas, P.J., Blais, J.M., Pauli, B.D., 2019. Using wood frog (Lithobates sylvaticus) tadpoles and semipermeable membrane devices to monitor polycyclic aromatic compounds in boreal wetlands in the oil sands region of northern Alberta, Canada. Chemosphere 214, 148-157. 10.1016/j.chemosphere.2018.09.034

National Academies, P., 2016. Spills of Diluted Bitumen From Pipelines : A Comparative Study of Environmental Fate, Effects, and Response. National Academies Press, Washington, DC. http://uml.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=e000xna&AN=1131269&site=ehost-live

Navarro-Martín, L., Lanctôt, C., Edge, C., Houlahan, J., Trudeau, V.L., 2012. Expression profiles of metamorphosis-related genes during natural transformations in tadpoles of wild Wood Frogs (Lithobates sylvaticus). Canadian Journal of Zoology 90, 1059-1071. 10.1139/z2012-074

OECD, 2009. Test No. 231: Amphibian Metamorphosis Assay. doi:https://doi.org/10.1787/9789264076242-en

OECD, 2015. Test No. 241: The Larval Amphibian Growth and Development Assay (LAGDA). doi:https://doi.org/10.1787/9789264242340-en

Palace, V., Peters, L., Berard-Brown, N., Smyth, P., Shanoff, T., Wiseman, G., Tomy, T., Havens, S., Dettman, H., Greer, C., Kajpust, H., Black, T., Hanson, M., Rodriguez-Gil, J.L., Taylor, E., 2021. Assessing Monitored Natural Recovery (MNR) for remediating crude oil spills in freshwater environments: The Freshwater Oil Spill Remediation Study (FOReSt) at the Experimental Lakes Area, Canada. International Oil Spill Conference Proceedings 2021. 10.7901/2169-3358-2021.1.667537

Patterson, S., 2019. The Toxic Effects of Oil Sands Contaminants on Developing Amphibians.

Patterson, S.A., Denton, D.T.J., Hasler, C.T., Blais, J.M., Hanson, M.L., Hollebone, B.P., Rodriguez-Gil, J.L., Langlois, V.S., Patey, G., Yang, Z., Orihel, D.M., 2022. Resilience of larval wood frogs (Rana sylvatica) to hydrocarbons and other compounds released from naturally weathered diluted bitumen in a boreal lake. Aquatic Toxicology 245, 106128. https://doi.org/10.1016/j.aquatox.2022.106128

Pedersen, T.L., 2022. patchwork: The Composer of Plots. https://CRAN.R-project.org/package=patchwork

Peterson, E.K., Buchwalter, D.B., Kerby, J.L., LeFauve, M.K., Varian-Ramos, C.W., Swaddle, J.P., 2017. Integrative behavioral ecotoxicology: bringing together fields to establish new insight to behavioral ecology, toxicology, and conservation. Current Zoology 63, 185-194. 10.1093/cz/zox010

Philibert, D.A., Philibert, C.P., Lewis, C., Tierney, K.B., 2016. Comparison of Diluted Bitumen (Dilbit) and Conventional Crude Oil Toxicity to Developing Zebrafish. Environmental Science & Technology 50, 6091-6098. 10.1021/acs.est.6b00949

Pollet, I., Bendell-Young, L.I., 2000. Amphibians as indicators of wetland quality in wetlands formed from oil sands effluent. Environmental Toxicology and Chemistry 19, 2589-2597. 10.1002/etc.5620191027

Polo-Cavia, N., Burraco, P., Gomez-Mestre, I., 2016. Low levels of chemical anthropogenic pollution may threaten amphibians by impairing predator recognition. Aquatic Toxicology 172, 30-35. https://doi.org/10.1016/j.aquatox.2015.12.019

Reeves, M.K., Perdue, M., Blakemore, G.D., Rinella, D.J., Holyoak, M., 2011. Twice as easy to catch? A toxicant and a predator cue cause additive reductions in larval amphibian activity. Ecosphere 2. 10.1890/es11-00046.1

Regnault, C., Usal, M., Veyrenc, S., Couturier, K., Batandier, C., Bulteau, A.-L., Lejon, D., Sapin, A., Combourieu, B., Chetiveaux, M., Le May, C., Lafond, T., Raveton, M., Reynaud, S., 2018. Unexpected metabolic disorders induced by endocrine disruptors in Xenopus tropicalis provide new lead for understanding amphibian decline. Proceedings of the National Academy of Sciences - PNAS 115, E4416-E4425. 10.1073/pnas.1721267115

Regnault, C., Willison, J., Veyrenc, S., Airieau, A., Méresse, P., Fortier, M., Fournier, M., Brousseau, P., Raveton, M., Reynaud, S., 2016. Metabolic and immune impairments induced by the endocrine disruptors benzo[a]pyrene and triclosan in Xenopus tropicalis. Chemosphere 155, 519-527. https://doi.org/10.1016/j.chemosphere.2016.04.047

Regnault, C., Worms, I.A.M., Oger-Desfeux, C., MelodeLima, C., Veyrenc, S., Bayle, M.-L., Combourieu, B., Bonin, A., Renaud, J., Raveton, M., Reynaud, S., 2014. Impaired liver function in Xenopus tropicalis exposed to benzo[a]pyrene: transcriptomic and metabolic evidence. BMC Genomics 15, 666. 10.1186/1471-2164-15-666

Rehman, K., Imran, A., Amin, I., Afzal, M., 2018. Inoculation with bacteria in floating treatment wetlands positively modulates the phytoremediation of oil field wastewater. Journal of Hazardous Materials 349, 242-251. https://doi.org/10.1016/j.jhazmat.2018.02.013

Rehman, K., Imran, A., Amin, I., Afzal, M., 2019. Enhancement of oil field-produced wastewater remediation by bacterially-augmented floating treatment wetlands. Chemosphere 217, 576-583. https://doi.org/10.1016/j.chemosphere.2018.11.041

Ritz, C., 2010. Toward a unified approach to dose-response modeling in ecotoxicology. Environmental toxicology and chemistry 29, 220-229. 10.1002/etc.7

Ritz, C., Baty, F., Streibig, J.C., Gerhard, D., 2016. Dose-Response Analysis Using R. PLOS ONE 10, e0146021. 10.1371/journal.pone.0146021

Robinson, D., Hayes, A., Couch, S., 2023. broom: Convert Statistical Objects into Tidy Tibbles. https://CRAN.R-project.org/package=broom

Robinson, S.A., Richardson, S.D., Dalton, R.L., Maisonneuve, F., Trudeau, V.L., Pauli, B.D., Lee-Jenkins, S.S.Y., 2017. Sublethal effects on wood frogs chronically exposed to environmentally relevant concentrations of two neonicotinoid insecticides. Environmental Toxicology and Chemistry 36, 1101-1109. 10.1002/etc.3739

Santos, R.G., Loh, W., Bannwart, A.C., Trevisan, O.V., 2014. An overview of heavy oil properties and its recovery and transportation methods. Brazilian Journal of Chemical Engineering 31, 571-590. 10.1590/0104-6632.20140313s00001853

Sloman, K.A., McNeil, P.L., 2012. Using physiology and behaviour to understand the responses of fish early life stages to toxicants. Journal of Fish Biology 81, 2175-2198. 10.1111/j.1095-8649.2012.03435.x

Sparling, D.W., 2010. Ecotoxicology of amphibians and reptiles, 2nd ed. ed. CRC Press/Taylor & Francis, Boca Raton [Fla.

Sparling, D.W., Linder, G., SETAC, Bishop, C.A., Krest, S., 2010. Ecotoxicology of Amphibians and Reptiles, Second Edition. Taylor & Francis. https://books.google.ca/books?id=YUhFAQAAIAAJ

Stanley, M.J., 2024. Engineered Floating Wetlands as a Secondary Oil Spill Remediation Strategy for Freshwater Shorelines. Department of Biosystems Engineering. The University of Manitoba, Winnipeg, Manitoba, Canada, p. 294.

Taylor, A.C., Kollros, J.J., 1946. Stages in the normal development of Rana pipiens larvae. The Anatomical Record 94, 7-23. 10.1002/ar.1090940103

Thambirajah, A.A., Koide, E.M., Imbery, J.J., Helbing, C.C., 2019. Contaminant and Environmental Influences on Thyroid Hormone Action in Amphibian Metamorphosis. Frontiers in endocrinology (Lausanne) 10, 276. 10.3389/fendo.2019.00276

Truter, J.C., van Wyk, J.H., Oberholster, P.J., Botha, A.M., Mokwena, L.M., 2017. an evaluation of the endocrine disruptive potential of crude oil water accommodated fractions and crude oil contaminated surface water to freshwater organisms using in vitro and in vivo approaches. Environmental Toxicology and Chemistry 36, 1330-1342. 10.1002/etc.3665

Urszán, T.J., Garamszegi, L.Z., Nagy, G., Hettyey, A., Török, J., Herczeg, G., 2015. No personality without experience? A test on Rana dalmatina tadpoles. Ecology and evolution 5, 5847-5856. 10.1002/ece3.1804

Whelan, A., Clark, J., Andrew, G., Michel, J., Benggio, B., 2014. Developing Cleanup Endpoints for Inland Oil Spills. International Oil Spill Conference Proceedings 2014, 1267-1280. 10.7901/2169-3358-2014.1.1267

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D.a., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the tidyverse. Journal of Open Source Software 4, 1686. 10.21105/joss.01686

Wickham, H., Seidel, D., 2022. scales: Scale Functions for Visualization. https://CRAN.R-project.org/package=scales

Xia, Z., Idowu, I., Halldorson, T., Lucas, A.-M., Stein, C., Kaur, M., Tomy, T., Marvin, C., Thomas, P.J., Hebert, C.E., Smith, R.A., Dwyer-Samuel, F., Provencher, J.F., Tomy, G.T., 2023. Microbead beating extraction of avian eggs for polycyclic aromatic compounds. Chemosphere 335, 139059. https://doi.org/10.1016/j.chemosphere.2023.139059

Xie, Y., 2014. knitr: A Comprehensive Tool for Reproducible Research in R. in: Stodden, V., Leisch, F., Peng, R.D. (Eds.). Implementing Reproducible Computational Research. Chapman and Hall/CRC.

Xie, Y., 2015. Dynamic Documents with R and knitr, 2nd ed. Chapman and Hall/CRC, Boca Raton, Florida. https://yihui.org/knitr/

Xie, Y., 2023. knitr: A General-Purpose Package for Dynamic Report Generation in R. https://yihui.org/knitr/

Xie, Y., Allaire, J.J., Grolemund, G., 2018. R Markdown: The Definitive Guide. Chapman and Hall/CRC, Boca Raton, Florida. https://bookdown.org/yihui/rmarkdown

Xie, Y., Dervieux, C., Riederer, E., 2020. R Markdown Cookbook. Chapman and Hall/CRC, Boca Raton, Florida. https://bookdown.org/yihui/rmarkdown-cookbook

Supplemental Tables

Supplemental Figures