Aquatic tri-trophic standardized microcosm TriCosm

Name Verena Riedl

Email verena.riedl@york.ac.uk

A standardized tri-trophic small-scale system (TriCosm) for the assessment of stressor induced effects on aquatic community dynamics

Verena Riedl, [a,b,*] Annika Agatz, [a] Rachel Benstead, [b] and Roman Ashauer, [a]

[a] University of York, Environment Department, Heslington, York, UK

[b] Fera Science Ltd., Centre for Chemical Safety and Stewardship, Sand Hutton, York, UK

[*] Address correspondence to verena.riedl@york.ac.uk

ABSTRACT

Chemical impacts on the environment are routinely assessed in single-species tests. They are employed to measure direct effects on non-target organisms but indirect effects on ecological interactions can only be detected in multi-species tests. Micro- and mesocosms are more complex and environmentally realistic, yet, they are less frequently used for environmental risk assessment because resource demand is high while repeatability and statistical power are often low. Test systems fulfilling regulatory needs (i.e. standardization, repeatability and replication) and the assessment of impacts on species interactions and indirect effects are lacking. Here we describe the development of the TriCosm, a repeatable aquatic multi-species test with three trophic levels and increased statistical power. High repeatability of community dynamics of three interacting aquatic populations (algae, *Ceriodaphnia*, *Hydra*) was found with an average coefficient of variation of 19.5% and the

ability to determine small effect sizes. The TriCosm combines benefits of both single-species tests (fulfillment of regulatory requirements) and complex multi-species tests (ecological relevance) and can be used, for instance at an intermediate tier in environmental risk assessment. Furthermore, comparatively quickly generated population and community toxicity data can be useful for the development and testing of mechanistic effect models.

KEYWORDS

ecological risk assessment, aquatic invertebrates, population-level effects, food chain, aquatic microcosm, multi-species testing

INTRODUCTION

The thorough assessment of environmental risks is essential for chemicals that could potentially be released into the environment. Agricultural pesticides, for instance, are used to enhance crop production but due to their toxic nature they may have negative effects on organisms other than the targeted species (Benton et al. 2007; Rockström et al. 2009; Beketov et al. 2013; Stehle and Schulz 2015).

The risks of pesticides to non-target organisms are routinely assessed in i) simple single-species tests at lower tiers and, if lower tier assessments raise concern, in ii) complex microcosms or mesocosms (European Food and Safety Authority (EFSA) 2013). However, systems that bridge the gap between the two alternatives to an intermediate level of complexity are lacking. The former require less effort and rapidly deliver large amounts of highly repeatable data on the performance of individual non-target organisms. The information obtained is, however, often of low ecological relevance as it is not directly relevant at the population and community level (Fleeger et al. 2003; Liebig et al. 2008). In

51 contrast, micro-/mesocosms are environmentally more realistic, yet, they are less frequently
52 used in environmental risk assessment. Unlike single-species tests, they are resource, time
53 and effort demanding. A variety of ecologically interacting factors can rapidly lead to
54 divergent system dynamics and increase the variance between replicates impeding the
55 understanding of dose-response relationships (Landis et al. 1997).
56 The importance of integrating environmental complexity into testing approaches has been
57 acknowledged as a priority for the assessment of chemical safety (Landis et al. 1997; Landis
58 2003; Bednarska et al. 2013; Scientific Committee on Health and Environmental Risks
59 (SCHER) et al. 2013). Chemical exposure could trigger indirect effects through interactions
60 with the environmental context such as the hydrological regime (Stampfli et al. 2013),
61 temperature (Moe et al. 2013), food quality (Campos et al. 2014) or other organisms (Del
62 Arco et al. 2015; Viaene et al. 2015); indirect effects have important implications for the
63 sensitivity of communities (Fleeger et al. 2003).
64 Organisms living in a contaminated environment may be pushed towards the boundaries
65 of their ecological niche and become more susceptible to additional stressors (Van Straalen
66 2003; Bednarska et al. 2013). Food chain processes, such as competition for food and altered
67 predation were shown to be particularly relevant to determine the magnitude of toxic effects
68 (Heugens et al. 2001; Bednarska et al. 2013; Kattwinkel et al. 2015). For instance, the no
69 observed effect concentration (NOEC) of the herbicide prometryn to ciliates was found to be
70 approximately 145 times lower in a bi-trophic microcosm compared to single-species tests.
71 The lower threshold was likely caused in response to an indirect and toxicant induced
72 reduction of food (Liebig et al. 2008). Intraspecific competition can also change the
73 sensitivity to pesticides (Foit, Kaske, and Liess 2012; Viaene et al. 2015) and indirectly
74 altered predation rates can lead to cascading effects on other trophic interactions and
75 ecosystem functions (Englert et al. 2012; Agatz et al. 2014; Viaene et al. 2015). Multi-species

testing using environmentally more relevant approaches, i.e. at the population and community level is clearly needed to assess indirect toxicant effects such as shifts in ecological interactions (Fleeger et al. 2003; Benton et al. 2007).

The necessity towards an inclusion of ecological interactions in chemical impact testing was described 10 years ago when a review on 14 years of pesticide studies in freshwater test systems was published (Relyea and Hoverman 2006). At the time, the authors found only 133 studies with at least two potentially interacting species of which only 17 studies focused on three trophic levels with producers, herbivores and carnivores.

Yet, microcosms that describe impacts on populations and/or communities in systems smaller than 10 L are rare (Metcalf et al. 1971; Daam and Van Den Brink 2007; Liebig et al. 2008; Englert et al. 2012; Foit, Kaske, and Liess 2012; Dolciotti et al. 2014; Del Arco et al. 2015; Viaene et al. 2015). Mostly they were used to focus on impacts on intra- or interspecific competition (one trophic level) (Foit, Kaske, and Liess 2012; Dolciotti et al. 2014; Del Arco et al. 2015; Viaene et al. 2015) or on consumer-resource relationships (two trophic levels) with herbivore-producer (Daam and Van Den Brink 2007) or predator-prey interactions (Barry and Davies 2004; Liebig et al. 2008; Englert et al. 2012). Tri-trophic systems are frequently used in terrestrial research, for example in plant-herbivore-parasite systems (Bredeson et al. 2015; Uhl et al. 2015) but few small test systems exist to assess direct and indirect impacts at the population and community level in the aquatic environment. Test formats include simulations of microbial detritus food chains (producer-consumer-decomposer (Fuma et al. 2000; Dawoud et al. 2017)) and producer-consumer communities with either invertebrate predator (Barry and Davies 2004) or vertebrate predator (Metcalf et al. 1971). Microbial tests were often conducted in culture flasks (250 ml) (Fuma et al. 2000; Dawoud et al. 2017) and small macroinvertebrate community tests were performed in

systems of few litres, for example in 10 L (Barry and Davies 2004) and 7 L systems (Metcalf et al. 1971).

Still, single-species systems appear convenient because they fulfill the regulatory needs for international standardization of test procedures, comparability of effect data, repeatability and replication (Liebig et al. 2008). Standardized and repeatable multispecies systems of intermediate complexity that bridge the simplicity of single species tests and the complexity of microcosms, yet fulfill regulatory requirements, are rare. To our knowledge, there is only one standardized microcosm (Taub 1989) available that falls into this category. The aquatic system was registered for pesticide testing (American Society for Testing of Materials (ASTM) 2011) and effects on two trophic levels covering ten primary producer and five primary consumer species can be assessed. It is, however, rarely used for standardized effect assessment, perhaps due to its relative complexity and the lack of mechanistic understanding of the interactions between species involved.

We developed a new test system with species interacting across three trophic levels and increased statistical power (i.e. standardization and low replicate variability).The system was designed to be cost-effective, rapid, repeatable with well understood population dynamics to i) allow the detection of small changes in population dynamics due to direct and indirect interactions, and ii) link observed effects to known system processes. Here we describe the standardized aquatic tri-trophic microcosm (hereafter TriCosm) focusing on system design and variability in the control treatment.

MATERIALS AND METHODS

Test organisms

The TriCosm comprises populations of the green alga *Pseudokirchneriella subcapitata*, the cladoceran *Ceriodaphnia dubia* and the cnidarian *Hydra viridissima.* This

dynamic food-chain is subject to fluctuating but predictable changes in food supply and intraspecific competition and is interconnected through consumer-resource relationships. The species were chosen based on their rapid life cycles and their sensitivity to toxicants. The green alga *P. subcapitata* and the cladoceran *C. dubia* are routinely used for tests in the regulatory risk assessment framework (Organization for Economic Co-operation and Development (OECD) 2004; OECD 2006; OECD 2012).

P. subcapitata stock cultures were obtained from the Culture Collection of Algae and Protozoa (CCAP, Scotland, UK) and used to initiate a culture line prior to each study and cultured in OECD media (OECD 2006). *C. dubia* were obtained from Unilever (Safety and Environmental Assurance Centre, Bedford, UK) and cultured as age specific cultures in moderately hard, synthetic freshwater (United States Environmental Protection Agency 2002). They were fed five times per week with a suspension of yeast, cerophyl® and trout-chow (~3.5 ml) and *P. subcapitata* (~11 x 10^7 cells/day) (United States Environmental Protection Agency 2002). *H. viridissima* were obtained from the Department of Evolutionary Zoology (University of Debrecen, Hungary), cultured in modified T82MV medium (modified after ASTM E1366-11 2011, Table SI 1, Table SI 2) and fed with newly hatched *Artemia salina* three times per week ad libitum. Both animal cultures were kept at 25 ± 1 °C and 12/12h light/dark cycle.

The TriCosm

TriCosms consist of Pyrex® crystallizing dishes (Sigma-Aldrich, UK) filled with 500 ml of T82MV medium (ASTM E1366-11 2011, Table SI 1, Table SI 2) that was determined as suitable for each species. The systems were covered with transparent watch glasses (diameter 125 mm; Sigma-Aldrich, UK) and positioned on an orbital laboratory shaker (Adolf Kuehner AG Switzerland, Type LS-W) set at 65 rpm throughout the test. The

150 experiments were set up for 21 days at 25 ± 1 °C, 12/12h light/dark, 1100 lux at the water
151 surface with cool white fluorescent light tubes 58.5 W (approx. 1.3 m above the test vessels).
152 TriCosms were started with *P. subcapitata* cells from a culture in logarithmic growth phase,
153 *C. dubia* neonates (< 24h age) from the third or fourth brood of cultured mothers and *H.*
154 *viridissima* without visible buds (≤ 2 d age). Water parameters (pH, dissolved O_2) and animal
155 and algal populations were monitored throughout the test duration two, three and five times
156 per week, respectively.
157
158 *Monitoring of the populations*
159 The systems were placed on an orbital shaker and slow shaking kept the algal cells
160 suspended. Only suspended algae were measured and no stirring was necessary before
161 sampling as preliminary studies showed significant correlation between suspended and total
162 algal concentrations (cells/ml) ($r_S = 0.98$, $p < 0.01$, $n = 90$, Figure SI 1). In-vivo fluorescence
163 activity of water subsamples (5 x 200 µl) was measured with a plate reader (Tecan® Infinite
164 200 PRO, settings Table SI 3) to determine the algal concentration (cells/ml).
165 *C. dubia* and *H. viridissima* were monitored with non-invasive methods to avoid
166 impacts on population dynamics and counted by eye three times per week. *C. dubia* were
167 visually grouped in two age-classes, juveniles and adults based on their dimensional
168 similarity with individuals in cultures aged younger or older than 4 d. All manual counts were
169 repeated until count differences did not exceed 20% of the lower value.
170
171 *Assessing a suitable community composition*
172 Tests with different setups were performed to optimize replicate variability, test
173 duration, addition times and densities for each species. A full factorial design for density and
174 timing was not feasible due to a too high number of possible combinations. Hence,

175 preliminary tests were performed to determine which algae-grazer combination in terms of
176 organism abundance would prevent both algal blooms and the death of grazers due to
177 starvation. No preliminary tests were done to determine the impact of *Hydra* predation on *C.*
178 *dubia* numbers prior to the test outlined in Table 1. Two organism densities and different
179 addition times were chosen based on preliminary testing and four different setups were
180 conducted simultaneously (Table 1). *C. dubia* were added on the same day as the green algae
181 in all experiments, except for setup 1 where grazers were added 1 day later to allow short
182 acclimation of the algae to test conditions. Dependent on food concentrations, *C. dubia*
183 matured later in setups 1, 2 than in setups 3, 4 hence *H. viridissima* were introduced to the
184 systems 5 and 4 days, respectively, after *C. dubia* were added. The predators were added only
185 once the grazers started reproducing to prevent variable numbers of *C. dubia* reproducers and
186 neonates early on in the systems that could lead to noticeable impacts on community
187 dynamics and replicate variability. Replicate numbers differed between setups 1, 2 and 3, 4
188 due to space constraints on the shaker platform.
189
190 *Validation of an optimal experimental setup*
191 The coefficient of variation (CV) was calculated as a standardized measure of
192 variance between replicates. It was expressed as a percentage and indicates the magnitude of
193 the standard deviation in comparison to the mean. Thus, if the inherent variability between
194 replicates is large compared to the size of the measured endpoint (e.g. animal abundance) a
195 significant treatment effect could only be detected if the response was very large (Sanderson
196 2002).
197 We computed the CVs for the algal concentration (cells/ml) and the total number of
198 *C. dubia* and *H. viridissima* on every sampling day. The values were then compared within
199 and between setups to monitor replicate variation over time and to determine an optimal setup

200 in terms of low variance. The experimental setup with the lowest replicate variability
201 (Experiment 1) was determined and repeated (Experiment 2) to assess the reproducibility of
202 system dynamics and replicate variability.
203
204 *Population dynamics and interactions*
205 In the interacting system, the intermediate trophic layer is directly affected by both
206 variations of food availability and predation strength, while indirect effects between the
207 bottom and top trophic level regulate a bottom up or top down controlled system. A
208 comparison of algal and grazer dynamics between systems where grazers are subject to i)
209 variations of food but not to predation and ii) a combination of food limitation and predation
210 can thus yield information on species interaction strength and whether the system is
211 controlled by bottom up or top-down effects. Hence, we performed additional tests with i)
212 only algae (Experiment 2A, n = 8) and ii) algae and grazers (Experiment 2B, n = 8). The
213 experiments were carried out simultaneously to Experiment 2 and according to the
214 experimental conditions and test setup used as for Experiments 1 and 2 (see Table **1** for
215 details). Experiments 2A and 2B were then compared to determine direct impacts of grazers
216 on algal dynamics and Experiments 2 and 2B were compared to assess i) indirect effects of
217 predators on algal growth and ii) direct effects of predators on *C. dubia* population
218 trajectories.
219
220 *Statistical analyses*
221 An a priori power analysis was performed to estimate minimum detectable response
222 sizes between control and treated TriCosm populations and increase the reliability and
223 transparency of the derived endpoints (EFSA 2013). The minimum detectable difference
224 (MDD), i.e. the size of a variation between sample averages required to be detected as

significantly different, is dependent on the chosen Type I error value α, the number of replicates employed and on the inherent variance such as replicate variability and/or sampling error (Brock et al. 2015). Here, we computed the MDD using the CVs assuming similar variance among controls and treatments. We hypothesized the use of 8 replicates and estimated the sensitivity of the TriCosm to reveal chemical effects for each population and at each sampling point. The MDD was calculated as described by Brain et al. (2005) from Sokal and Rohlf (1995):

$$\mathrm{MDD} = \frac{\sqrt{2}\left(\mathrm{t}_{\alpha,\nu} + \mathrm{t}_{\beta,\nu}\right) \mathrm{x\ CV}}{\sqrt{n}}$$

where $\mathrm{t}_{\alpha,\nu}$ and $\mathrm{t}_{\beta,\nu}$ are the t-values for α and β set to 0.05 and 0.2, respectively, for a confidence level of 95% and a power of 80% at ν degrees of freedom. CV is the coefficient of variation and n is the number of replicates used. The degrees of freedom were computed as ν = k (n-1) and the number of groups k was set to 2, e.g. to compare each treatment to the control. The calculated MDDs were compared to MDD classes as proposed by the European Food Safety Authority (EFSA) that grouped MDD sizes into five classes and described the likely ability of effect detection (Class 0: MDD > 100% = no effect detection, Class I: MDD 90 – 100% = only large effects, Class II: MDD 70 – 90% = large to medium effects, Class III: MDDs 50 – 70% = medium effects, Class IV: MDD < 50% small effects) (EFSA 2013).

To assess species interactions between system components, population dynamics were compared graphically between experiments and significant differences were assumed where 95% confidence intervals did not overlap.

RESULTS

Optimizing the experimental setup

The community dynamics (

Figure 1) and the coefficients of variation differed among the four setups (Figure 2) and over time (Figure SI 2). In general, the algal concentration (cells/ml) peaks were followed by *C. dubia* abundance peaks and a constant increase of *H. viridissima* populations. The highest *C. dubia* peak 172 (± 10) individuals (mean ± 95% confidence interval range) was found in Setup 1 on day 14 following an algal peak on day 6 when an average of 6.86 (± 0.64) x 10^5 cells/ml was measured (

Figure 1A).The highest algal peak of 11.51 (± 5.59) x 10^5 cells/ml appeared in Setup 3 on day 11 with increasing variance in terms of organism numbers between replicates over time (

Figure 1C). *H. viridissima* populations showed steady growth during the test duration and increased in numbers by an average of 42 (± 6), 45 (± 8), 72 (± 18) and 54 (± 15) individuals in Setup 1, 2, 3 and 4, respectively (Figure 1A - D). Final counts differed due to different addition numbers, timings and food availability (*C. dubia* abundances), however, *Hydra* populations showed the smallest replicate variance when compared to algae and *C. dubia* (Figure 2). The CVs of all test variables in each setup increased over time (Figure SI 2) with setup 3 showing the highest replicate variability (with the exception of *H. viridissima*) and setup 1 showing the lowest average CV (with the exception of *H. viridissima)* (Figure 2). High CVs observed in setup 2 - 4 indicated reduced ability to detect treatment related system alterations. Therefore we selected setup 1 (Experiment 1) as most appropriate setup procedure (Protocol SI 4) and repeated the test (Experiment 2) to evaluate the repeatability of the system.

Validation of the test setup

The population dynamics of experiment 1 and the repeated experiment 2 were similar (Figure 3). The algal populations peaked on day 6 and day 5 with average algal concentrations of 6.86 (± 0.64) x 10^5 and 8.83 (± 0.90) x 10^5 cells/ml in experiments 1 and 2, respectively (Figure 3A). The highest *C. dubia* abundance was measured 8 days after the algal peaks in both cases. Grazer numbers declined 1 day earlier in experiment 2 and individual counts were lower due to smaller juvenile numbers of 147 (± 10) and 97 (± 24) juveniles in experiment 1 and 2, respectively (Figure 3B). The dynamics of *H. viridissima* populations were similar between experiments (Figure 3C) but experiment 1 showed a slightly steeper population increase with a larger final population of 45 (± 6) and 33 (± 9) individuals in experiment 1 and 2, respectively. Due to slightly different sampling frequencies, we computed 15, 11 and 8 CV values in experiment 1 and 14, 9 and 7 CVs in experiment 2 for algal concentrations (cells/ml), *C. dubia* and *H. viridissima*, respectively. As observed for population dynamics, replicate variance was similar between populations in both experiments (Figure SI 3). The CVs of algal concentrations (cells/ml) increased by day 7, 14 and 21 to an average of 9, 26 and 26% in experiment 1 and 10, 25 and 47% in experiment 2, respectively. The CVs calculated for *C. dubia* populations increased from 9 to 15 and 33% in experiment 1 and from 4 to 23 and 24% in experiment 2. *H. viridissima* were added on day 6, so the replicate variability was 18, 18% and 17, 29% by day 14 and 21 in experiment 1 and 2, respectively.

Population dynamics and interactions

Significant reductions of algal concentrations (cells/ml) by 33.4% were found on the first sampling day after *C. dubia* addition, on day 4. An average algal concentration of 9.18 (± 0.48) x 10^5 cells/ml was found in Experiment 2A (only algae), while grazed algae in Experiment 2B (algae and grazers) reached an abundance of 6.11 (± 0.43) x 10^5 cells/ml by

298 day 4 and stayed significantly lower throughout the experimental duration (Figure 4A).
299 Grazed algal concentrations in systems with and without predator (Experiments 2 and 2B,
300 respectively) showed similar trajectories until day 14. After day 14, algal concentrations in
301 Experiment 2B stayed moderately constant with an average of 2.12 (± 0.21) x 10^5 cells/ml
302 until day 20. On the contrary, algal concentrations in Experiment 2 (grazers and predators)
303 increased to 7.63 (± 0.37) x 10^5 cells/ml by day 20 exceeding average algal abundances of
304 Experiment 2B by 75.0% (Figure 4B).
305 Grazer population dynamics were similar in Experiment 2 and 2B until day 6 when
306 predators were added to Experiment 2 (Figure 4C). Population numbers peaked in both
307 experiments on day 13 but steeper population growth curves in Experiment 2B lead to an on
308 average 39.1% larger peaking population of 210 (± 21) individuals when compared to 128 (±
309 25) individuals in Experiment 2. A similarly steeper population decrease in experiment 2B
310 resulted in similar total *C. dubia* counts of 36 (± 9) and 37 (± 10) individuals in systems
311 without and with predators, respectively, by day 20. Population dynamics of total grazer
312 numbers largely reflected the trajectories of *C. dubia* juveniles that rapidly increased until
313 day 13 to 175 (± 23) and 107 (± 23) individuals constituting 88.3% and 90.1% of the total *C.*
314 *dubia* population in experiments 2B and 2, respectively. By day 20, juvenile numbers
315 dropped to 6 (± 4) and 15 (± 6) individuals while adult grazers showed a moderate but
316 constant increase throughout the test and constituted 84.9% and 59.7% of the total *C. dubia*
317 populations in experiments without and with predators, respectively (Figure 4D).
318
319 *Statistical analyses*
320 Minimum detectable differences (MDDs) of hypothetical TriCosm exposures were
321 calculated according to control variance and were similar between experiments 1 and 2. The
322 TriCosm became less sensitive over time as replicate variation and MDDs increased. When

323 variances between controls and treatments are similar, the TriCosm is estimated to be
324 sufficiently sensitive to identify differences of 12% (± 4), 36% (± 7) and 50% (± 17) for *P.*
325 *subcapitata* and 9% (± 7), 31% (± 4) and 38% (± 19) for *C. dubia* populations in week 1
326 (day 1 – 6), week 2 (day 7 – 13) and week 3 (day 14 – 21), respectively (averaged MDDs
327 between Experiment 1 and 2). Averaged MDDs for *H. viridissima* were 25% (± 14) and 35%
328 (± 7) in week 2 and 3, respectively (Figure 3). The MDDs for critical endpoints in the
329 TriCosm can be assigned to MDD classes III (50 – 70%) and IV (< 50%), indicating the
330 ability to determine small and medium sized effects, respectively (EFSA 2013).
331
332 **DISCUSSION**
333 The assessment of chemical effects with single-species tests fulfills regulatory
334 requirements; however, primary goals of protecting populations and ecosystems might not be
335 adequately addressed. That is because information obtained at the individual level is often not
336 ecologically relevant since there are neither directly proportionate relationships between
337 direct and indirect effects nor amongst responses at the individual, population and community
338 level. An understanding of impacts on interactions in ecologically relevant test settings is thus
339 critical and a priority for chemical safety assessment as unexpected shifts in community
340 profiles cannot be predicted in single-species tests (Fleeger et al. 2003; Benton et al. 2007;
341 Liebig et al. 2008; SCHER et al. 2013).
342 We designed the TriCosm as a rapidly cycling, tri-trophic system with a producer-
343 herbivore-carnivore community of small size for the purpose of quick detection of chemical
344 impacts on species interactions. Our system is comparatively smaller (0.5 L) than many other
345 multi-trophic systems (Metcalf et al. 1971; Daam and Van Den Brink 2007; Foit, Kaske,
346 Wahrendorf, et al. 2012; Dolciotti et al. 2014; Del Arco et al. 2015) and all system
347 components exhibit rapid generation times so that treatment effects can be measured on

several generations and at different life stages during short test durations (21 days compared to 80 days (Metcalf et al. 1971) and 33 days (Barry and Davies 2004) in other tritrophic macroinvertebrate communities). Also the predator *Hydra* is a rapid reproducer with generation times of only three days under favourable conditions (Habetha et al. 2003). Chemical impacts on population dynamics can thus be detected not only at the producer-consumer level but also at a higher trophic level. The choice of a small and rapidly reproducing predator has further the advantage that it can be added at an early experimental stage (day 6) when compared to vertebrate predators that are often introduced shortly before test termination as they quickly consume remaining invertebrate preys (Metcalf et al. 1971; Harrass and Taub 1985).

All multi-species systems have ecologically interacting components that are not independent in statistical terms as they constantly adapt to changing conditions in a dynamic environment. In fact, it has been frequently reported that even though communities are set up identically as replicates, minor variations at the beginning and/or throughout the experiments (e.g. starting conditions or uneven sample removal) can quickly lead to the development of unique properties in each replicate (Landis et al. 1997; Sanderson 2002; Van Straalen 2003). Indeed, different population dynamics and replicate variability were observed in four different TriCosm setups and indicated strong sensitivity to starting conditions and interaction strength. The statistical quality (in terms of interpretability, reproducibility and replicability) of environmentally more realistic data obtained in multi-species tests is thus often reason for concern in the registration procedure of pesticides (Sanderson 2002).

The repeatability and reproducibility of the TriCosm were thus given major consideration during test development. Initial properties and sampling techniques were adjusted and confirmed as optimized when experiments conducted at different times showed low coefficients of variation (CVs) and high reproducibility of system dynamics.

373 Desynchronized population dynamics were observed between experiments that can be
374 attributed to random fluctuations in test conditions (e.g. quality of the animals) and could
375 occur even if procedures are standardized. For these reasons we assessed the repeatability by
376 comparing CVs and not the total organism abundances. Nonetheless, a comparison of total
377 abundances or derived variables (e.g. population growth rates) is also appropriate when
378 chemical effects are assessed since differences between population trajectories are most
379 likely and primarily due to chemical impacts rather than fluctuations of test conditions.
380 When the TriCosm is used for chemical effect assessment, two factors of major
381 importance are i) the presence of interactions rather than the exact timing when these occur
382 and ii) low CVs so that treatment responses can be interpreted with greater certainty and
383 distinguished from unexplained sample variability (Sanderson 2002).
384 The ability to detect significant effects does depend on the magnitude of an effect but
385 also on the ability of the test system to detect responses and that is in turn determined by the
386 inherent variance among replicates. Test variables with coefficients of variation (CV) in the
387 range of up to 30% have been theorized as acceptable and manageable in terms of practicality
388 and costs (Kraufvelin 1998). According to a review (Sanderson 2002) that analyzed two
389 decades of pesticide studies with micro/mesocosms, the values of CVs appear to be generally
390 higher. The author reported an average of 45% (32% in smaller and less realistic indoor
391 systems) with larger values in studies where animals were involved and an average use of 3.5
392 replicates. The average CV of 19.5% measured in the tri-trophic system on the contrary
393 showed smaller variance and was determined with a higher number of replicates (n = 8).
394 The CVs were further used for the calculation of theoretically detectable minimum
395 differences (MDDs) between controls and treatments under the assumption of similar
396 variances. It is to be mentioned, however, that the variance could increase, decrease and/or
397 remain similar in treated systems (Kraufvelin 1998; Sanderson 2002). A modification of the

number of replicates, groups or treatments, though, can decrease MDDs and allow the detection of desired effect sizes. Due to often large variability in micro-/mesocosms, EFSA may still regard endpoints with MDD classes I and II (70 – 100%) relevant but considers the exceeding of class II ideal (i.e. MDDs < 70%) (EFSA 2013). Most projected MDDs of critical endpoints in the aquatic system correspond to effect class IV (i.e. < 50%) (with exception of algae and grazers in week 3) and confirm the ability to reveal small toxicant induced effect sizes (EFSA 2013), distinguishing the TriCosm from other multi-trophic systems.

As expected, variations of population trajectories were observed as a result of interactions with other system components. Algal concentrations (cells/ml) and predation both directly impacted on the middle trophic layer while they indirectly impacted on the top and the bottom level, respectively. An initially small grazing pressure of juvenile *C. dubia* allowed algal populations to grow exponentially which in turn favored the development of grazer populations. As a consequence of an increasing grazing pressure by maturing and reproducing *C. dubia,* the algal concentrations dropped, yet the grazer population numbers further increased for approximately one week after food availability became limiting. The continuing population growth is attributable to a rise of juvenile numbers as adult *C. dubia* most likely matured eggs and stored energy before algal concentrations decreased. Peaking *C. dubia* populations thus coincided with lows of food availability and caused the decrease of grazer numbers. Algae stabilized and remained at relatively constant levels as concentrations were most likely too low to be further reduced if maximum grazer filtering rates were reached. Grazer population numbers decreased due to food shortage and independently of whether predators were present or not. While predation did not cause *C. dubia* populations to crash, it directly reduced grazer numbers, intraspecific competition among them and indirectly favored algal populations to recover. An increase of algal concentrations after

423 grazing release was, however, not observed in Experiment 1 where grazer populations
424 reached larger abundances but decreased later and might be due to a different quality of
425 neonates used to start the experiments. Algal populations in Experiment 1 were thus subject
426 to a higher and prolonged grazing pressure impeding the recovery of algal abundances within
427 the experimental duration. An indirect effect after grazing release by *Hydra* could, however,
428 likely be expected if the test duration was prolonged. Bottom up and top down processes are
429 thus both likely regulating population dynamics in the TriCosm. When the system is exposed
430 to chemicals it will thus depend on the mode of action of the toxicant impacting on one or
431 more trophic levels leading to direct, indirect or both effects on the trajectories of interacting
432 populations.

433 The focus during system development was not on achieving a steady state community
434 and impacts on resilience cannot be assessed, neither can system shifts be detected.
435 Nonetheless, it can indicate the recovery potential of species after stressor removal and detect
436 small changes in interactions as the system moves through a single cycle of the middle
437 trophic layer. Ecological impacts of toxicants rapidly propagate in an interacting system and
438 the grazer level is directly influenced by variations in food availability and predation.
439 Toxicant impacts on the population dynamics of this critical and key trophic layer will
440 therefore yield important information on the ecological relevance and protectiveness of data
441 obtained in single-species tests. Population responses to combined stressor exposures, e.g. to
442 toxicants, predation and/or food fluctuations, could be used to facilitate both the development
443 and the testing of mechanistic effect models. Measured community responses in terms of
444 individual abundance changes and population trajectories could be employed for the
445 calibration and parameter fitting of ecological models. In turn, chemical effects on
446 interactions within a simple freshwater community can be measured and quantified in the

TriCosm and provide empirical benchmarking to estimate and test model prediction accuracy and power.

There is no doubt that the complexity of the TriCosm community is low when compared to natural systems. But besides offering higher statistical power when compared to larger and / or more complex microcosms, the impacts on system processes can be quantified as interactions change. This makes it possible to assess the effects of environmental contaminants on i) species interactions, ii) indirect effects and iii) at the population and community level. An understanding of which and to what extent processes are affected may also give insights into responses of more complex systems (Benton et al. 2007; Daam and Van Den Brink 2007; Boonstra et al. 2011) .

CONCLUSION

The TriCosm is a novel aquatic test system and could be a tool to address shifts in ecological interactions. It suggests that a cost-effective approach of chemical environmental safety testing with more ecological relevance whilst being statistically powerful is feasible. It can provide important insights into chemical safety in multi-trophic systems and facilitate the development and testing of mechanistic effect models for environmental risk assessment. Even so, a careful examination of the replicability of the TriCosm both within and between laboratories with and without chemical exposure is needed.

Supplemental Data — The Supplemental Data are available on the Wiley Online Library at DOI: 10.1002/etc.xxxx

Acknowledgment — This study was financed by The Environment Department, University of York, York, UK, and the Centre for Chemical Safety and Stewardship (CCSS), Fera Science

Ltd., York, UK. The authors thank J. Turton and T. McGowan (CCSS) for technical support and two anonymous reviewers for their comments helping improve the manuscript. The authors have no conflicts of interest.

Data availability — Data and calculation tools are available from the corresponding author (Verena.riedl@york.ac.uk).

REFERENCES

Agatz A, Ashauer R, Brown CD. 2014. Imidacloprid perturbs feeding of *Gammarus pulex* at environmentally relevant concentrations. Environ. Toxicol. Chem. 33:648–653. doi:10.1002/etc.2480.

American Society for Testing of Materials E1366-11. 2011. Standard Practice for Standardized Aquatic Microcosms : Fresh Water. West Conshohocken, PA: ASTM International.

Del Arco AI, Parra G, Rico A, Van den Brink PJ. 2015. Effects of intra- and interspecific competition on the sensitivity of aquatic macroinvertebrates to carbendazim. Ecotoxicol. Environ. Saf. 120:27–34. doi:10.1016/j.ecoenv.2015.05.001.

Barry MJ, Davies W. 2004. Effects of invertebrate predators and a pesticide on temporary pond microcosms used for aquatic toxicity testing. Environ. Pollut. 131:25–34. doi:10.1016/j.envpol.2004.02.023.

Bednarska AJ, Jevtić DM, Laskowski R. 2013. More ecological ERA: incorporating natural environmental factors and animal behavior. Integr. Environ. Assess. Manag. 9999:1–8. doi:10.1002/ieam.1444.

Beketov MA, Kefford BJ, Schäfer RB, Liess M. 2013. Pesticides reduce regional biodiversity of stream invertebrates. Proc. Natl. Acad. Sci. U. S. A. 110:11039–43. doi:10.1073/pnas.1305618110.

Benton TG, Solan M, Travis JM, Sait SM. 2007. Microcosm experiments can inform global ecological problems. Trends Ecol Evol 22:516–521. doi:10.1016/j.tree.2007.08.003.

Boonstra H, Reichman EP, Van Den Brink PJ. 2011. Effects of the veterinary pharmaceutical ivermectin in indoor aquatic microcosms. Arch. Environ. Contam. Toxicol. 60:77–89. doi:10.1007/s00244-010-9526-1.

Brain RA, Bestari KJ, Sanderson H, Hanson ML, Wilson CJ, Johnson DJ, Sibley PK, Solomon KR. 2005. Aquatic microcosm assessment of the effects of tylosin on *Lemna gibba* and *Myriophyllum spicatum*. 133:389–401. doi:10.1016/j.envpol.2004.07.005.

Bredeson MM, Reese RN, Lundgren JG. 2015. The effects of insecticide dose and herbivore density on tri-trophic effects of thiamethoxam in a system involving wheat, aphids, and ladybeetles. Crop Prot. 69:70–76. doi:10.1016/j.cropro.2014.12.010.

Brock TCM, Hammers-Wirtz M, Hommen U, Preuss TG, Ratte HT, Roessink I, Strauss T, Van den Brink PJ. 2015. The minimum detectable difference (MDD) and the interpretation of treatment-related effects of pesticides in experimental ecosystems. Environ. Sci. Pollut. Res. Int. 22:1160–1174. doi:10.1007/s11356-014-3398-2.

Campos D, Alves A, Lemos MFL, Correia A, Soares AMVM, Pestana JLT. 2014. Effects of cadmium and resource quality on freshwater detritus processing chains: A microcosm approach with two insect species. Ecotoxicology 23:830–839.

Daam M a., Van Den Brink PJ. 2007. Effects of chlorpyrifos, carbendazim, and linuron on the ecology of a small indoor aquatic microcosm. Arch. Environ. Contam. Toxicol. 53:22–35.

518 doi:10.1007/s00244-006-0001-y.

519 Dawoud M, Bundschuh M, Goedkoop W, McKie BG. 2017. Interactive effects of an
520 insecticide and a fungicide on different organism groups and ecosystem functioning in a
521 stream detrital food web. Aquat. Toxicol. 186:215–221. doi:10.1016/j.aquatox.2017.03.008.

522 Dolciotti I, Foit K, Herkelrath A, Liess M. 2014. Competition impedes the recovery of
523 *Daphnia magna* from repeated insecticide pulses. Aquat Toxicol 147:26–31.
524 doi:10.1016/j.aquatox.2013.12.002.

525 Englert D, Bundschuh M, Schulz R. 2012. Thiacloprid affects trophic interaction between
526 gammarids and mayflies. Env. Pollut 167:41–46. doi:10.1016/j.envpol.2012.03.024.

527 European Food and Safety Authority (EFSA) Panel on Plant Protection Products and their
528 Residues. 2013. Guidance on tiered risk assessment for plant protection products for aquatic
529 organisms in edge-of-field surface waters. EFSA J. 11:3290. doi:10.2903/j.efsa.2013.3290.

530 Fleeger JW, Carman KR, Nisbet RM. 2003. Indirect effects of contaminants in aquatic
531 ecosystems. Sci. Total Environ. 317:207–233. doi:10.1016/s0048-9697(03)00141-4.

532 Foit K, Kaske O, Liess M. 2012. Competition increases toxicant sensitivity and delays the
533 recovery of two interacting populations. Aquat. Toxicol. 106–107:25–31.
534 doi:10.1016/j.aquatox.2011.09.012.

535 Foit K, Kaske O, Wahrendorf D-S, Duquesne S, Liess M. 2012. Automated Nanocosm test
536 system to assess the effects of stressors on two interacting populations. Aquat Toxicol
537 109:243–249. doi:10.1016/j.aquatox.2011.09.013.

538 Fuma S, Takeda H, Miyamoto K, Yanagisawa K, Inoue Y, Ishii N, Sugai K, Ishii C,
539 Kawabata Z. 2000. Simple aquatic microcosm for ecotoxicity screening at the community

level. Bull. Environ. Contam. Toxicol. 65:699–706. doi:10.1007/s001280000180.

Habetha M, Anton-Erxleben F, Neumann K, Bosch TC. 2003. The Hydra viridis/Chlorella symbiosis. Growth and sexual differentiation in polyps without symbionts. Zool. 106:101–108. doi:10.1078/0944-2006-00104.

Harrass MC, Taub FB. 1985. Effects of small fish predation on microcosm community bioassays. Spec.tech.Publ.Am.Soc.Test.Mat. 854:117–133.

Heugens EHW, Hendriks AJ, Dekker T, Straalen NM van, Admiraal W. 2001. A Review of the Effects of Multiple Stressors on Aquatic Organisms and Analysis of Uncertainty Factors for Use in Risk Assessment. Crit. Rev. Toxicol. 31:247–284. doi:10.1080/20014091111695.

Kattwinkel M, Liess M, Arena M, Bopp S, Streissl F, Römbke J. 2015. Recovery of aquatic and terrestrial populations in the context of European pesticide risk assessment. Environ. Rev. 23:382–394. doi:10.1139/er-2015-0013.

Kraufvelin P. 1998. Model ecosystem replicability challenged by the "'soft'" reality of a hard bottom mesocosm. J. Exp. Mar. Bio. Ecol. 222:247–267.

Landis WG. 2003. Twenty years before and hence; Ecological risk assessment at multiple scales with multiple stressors and multiple endpoints. Hum. Ecol. Risk Assess. 9:1317–1326. doi:Doi 10.1080/10807030390248500.

Landis WG, Matthews RA, Matthews GB. 1997. Design and analysis of multispecies toxicity tests for pesticide registration. Ecol. Appl. 7:1111–1116. doi:10.1890/1051-0761(1997)007[1111:DAAOMT]2.0.CO;2.

Liebig M, Schmidt G, Bontje D, Kooi BW, Streck G, Traunspurger W, Knacker T. 2008. Direct and indirect effects of pollutants on algae and algivorous ciliates in an aquatic indoor

562 microcosm. Aquat. Toxicol. 88:102–110. doi:10.1016/j.aquatox.2008.03.010.

563 Metcalf RL, Gurcharan KS, Inder PK. 1971. Model Ecosystem for the Evaluation of
564 Pesticide Biodegradability and Ecological Magnification. Environ. Sci. Technol. 5:709–713.

565 Moe SJ, De Schamphelaere K, Clements WH, Sorensen MT, Van den Brink PJ, Liess M.
566 2013. Combined and interactive effects of global climate change and toxicants on populations
567 and communities. Environ. Toxicol. Chem. 32:49–61. doi:10.1002/etc.2045.

568 Organisation for Economic Co-operation and Development (OECD). 2004. OECD
569 Guidelines for the Testing of Chemicals, Guideline 202: Daphnia sp. Acute Immobilisation
570 Test. OECD Publishing, France OECD Guidelines for the Testing of Chemicals, Section 2.

571 Organisation for Economic Co-operation and Development (OECD). 2006. OECD
572 Guidelines for the Testing of Chemicals, Guideline 201: Freshwater Alga and Cyanobacteria,
573 Growth Inhibition Test. OECD Publishing, France.

574 Organisation for Economic Co-operation and Development (OECD). 2012. OECD
575 Guidelines for the Testing of Chemicals, Guideline 211: Daphnia magna Reproduction Test.
576 OECD Publishing, France.

577 Relyea R, Hoverman J. 2006. Assessing the ecology in ecotoxicology: A review and
578 synthesis in freshwater systems. Ecol. Lett. 9:1157–1171. doi:10.1111/j.1461-
579 0248.2006.00966.x.

580 Rockström J, Steffen W, Noone K, Persson Å, Chapin FS, Lambin E, Lenton TM, Scheffer
581 M, Folke C, Schellnhuber HJ, et al. 2009. Planetary boundaries: Exploring the safe operating
582 space for humanity. Ecol. Soc. 14:32.

583 Sanderson H. 2002. Replicability of Micro/Mesocosms. Environ. Sci. Pollut. Res. Int. 9:429–

435.

Scientific Committee on Health and Environmental Risks, Scientific Committee on Emerging and Newly Identified Health Risks, Scientific Committee on Consumer Safety. 2013. Addressing the New Challenges for Risk Assessment. :1–157. doi:10.2772/37863.

Sokal RR, Rohlf JF. 2012. Biometry: the principles of statistics in biological research. 4th editio. W. H. Freeman and Company New York.

Stampfli NC, Knillmann S, Liess M, Noskov Y a., Schäfer RB, Beketov M a. 2013. Two stressors and a community - Effects of hydrological disturbance and a toxicant on freshwater zooplankton. Aquat. Toxicol. 127:9–20. doi:10.1016/j.aquatox.2012.09.003.

Stehle S, Schulz R. 2015. Agricultural insecticides threaten surface waters at the global scale. Proc. Natl. Acad. Sci. 112:5750–5755. doi:10.1073/pnas.1500232112.

Van Straalen NM. 2003. Ecotoxicology becomes stress ecology. Environ. Sci. Technol. 37:324A–330A. doi:10.1021/es0325720.

Taub F. 1989. Standardized Aquatic Microcosms. Env. Sci Technol 23:1064–1066. doi:10.1021/es00067a601.

Uhl P, Bucher R, Schäfer RB, Entling MH. 2015. Sublethal effects of imidacloprid on interactions in a tritrophic system of non-target species. Chemosphere 132:152–158. doi:10.1016/j.chemosphere.2015.03.027. [accessed 2017 Aug 15]. http://linkinghub.elsevier.com/retrieve/pii/S0045653515002374.

United States Environmental Protection Agency. 2002. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms. Washington, DC EPA-821-R-02-013.

606 Viaene KPJ, De Laender F, Rico A, Van den Brink PJ, Di Guardo A, Morselli M, Janssen
607 CR. 2015. Species interactions and chemical stress: Combined effects of intraspecific and
608 interspecific interactions and pyrene on Daphnia magna population dynamics. Environ.
609 Toxicol. Chem. 34:1751–1759. doi:10.1002/etc.2973.

610

611

612 **TABLES AND FIGURES**

613 Table 1. The TriCosm community composition at the beginning of four test setups

	P. subcapitata		*C. dubia*		*H. viridissima*		
	Cells/ml	Day	Individuals	Day	Individuals	Day	Replicates
Setup 1	2×10^4	0	10	1	3	6	8
Setup 2	2×10^4	0	10	0	6	5	8
Setup 3	4×10^4	0	10	0	6	4	7
Setup 4	4×10^4	0	20	0	6	4	7

614

615

616 Figure 1 Algal concentrations (cells/ml) and total number of *C. dubia* and *H. viridissima* over
617 21 days. Shown are means ± 95 % confidence intervals in four test setups (A – D) (see Table
618 1 for details).

619

620 Figure 2 Coefficients of variation (%) of algal concentrations (cells/ml), total abundance of
621 *C. dubia* and *H. viridissima* at each sampling event. Black horizontal lines indicate 95 %
622 confidence intervals in setups 1 - 4.

623

Figure 3 Abundance of (A) *P. subcapitata*, (B) *C. dubia* and (C) *H. viridissima* at each sampling point over 21 days. Shown are means ± 95% confidence intervals and minimum detectable differences (% MDD) below the x- axis of experiment 1 and experiment 2 (Exp 1, Exp 2).

Figure 4 Algal population trajectories compared between (A) ungrazed (green circles) and grazed (without predation, blue diamonds) systems and (B) grazed systems with (red stars) and without predation (blue diamonds). Population dynamics of *C.dubia* with (red stars) and without (blue diamonds) predation as (C) total individual number and (D) juveniles (continuous line) and adults (dotted line).

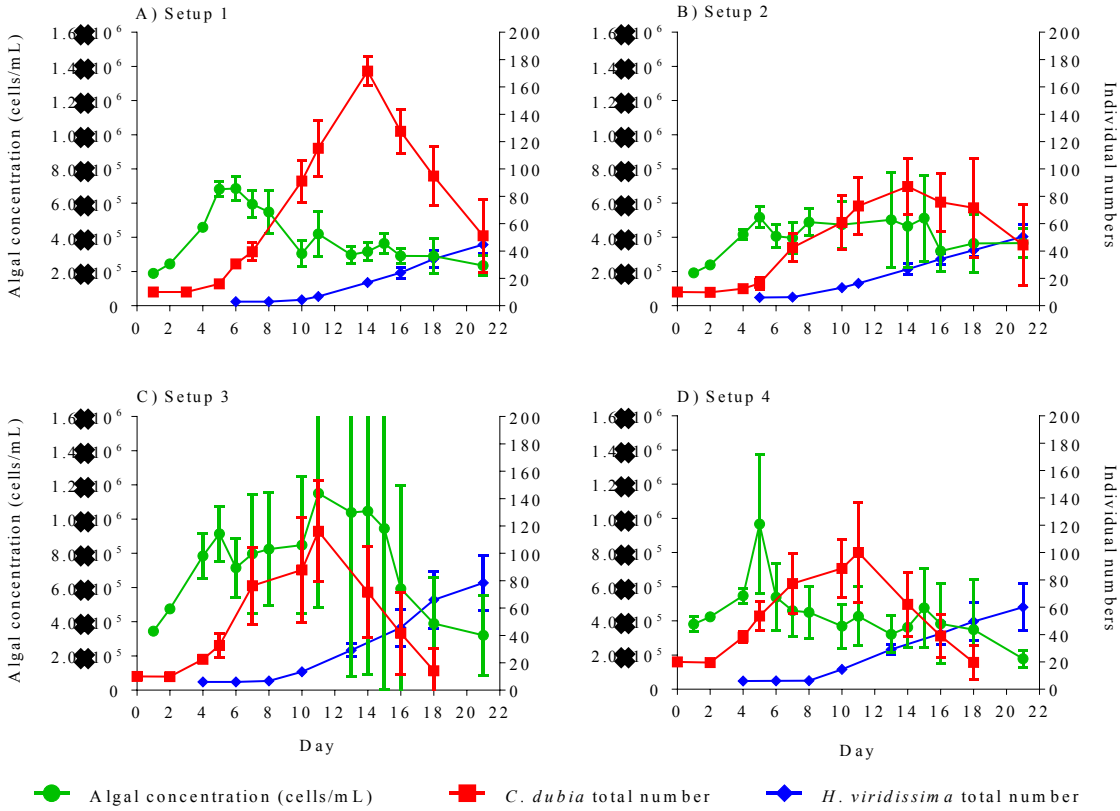
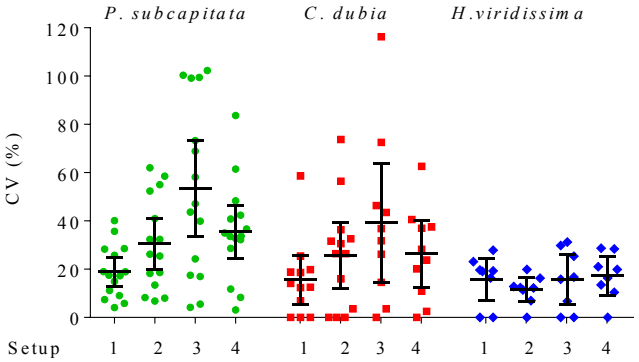


Figure 5

637

638

639

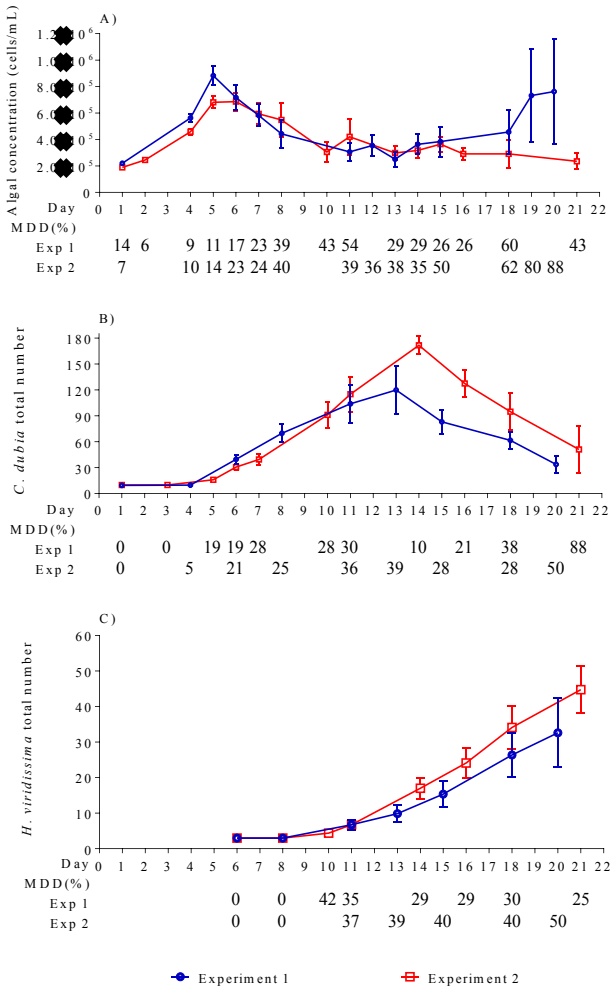


640

641 Figure 2

642

643



644

645 Figure 3

646

647

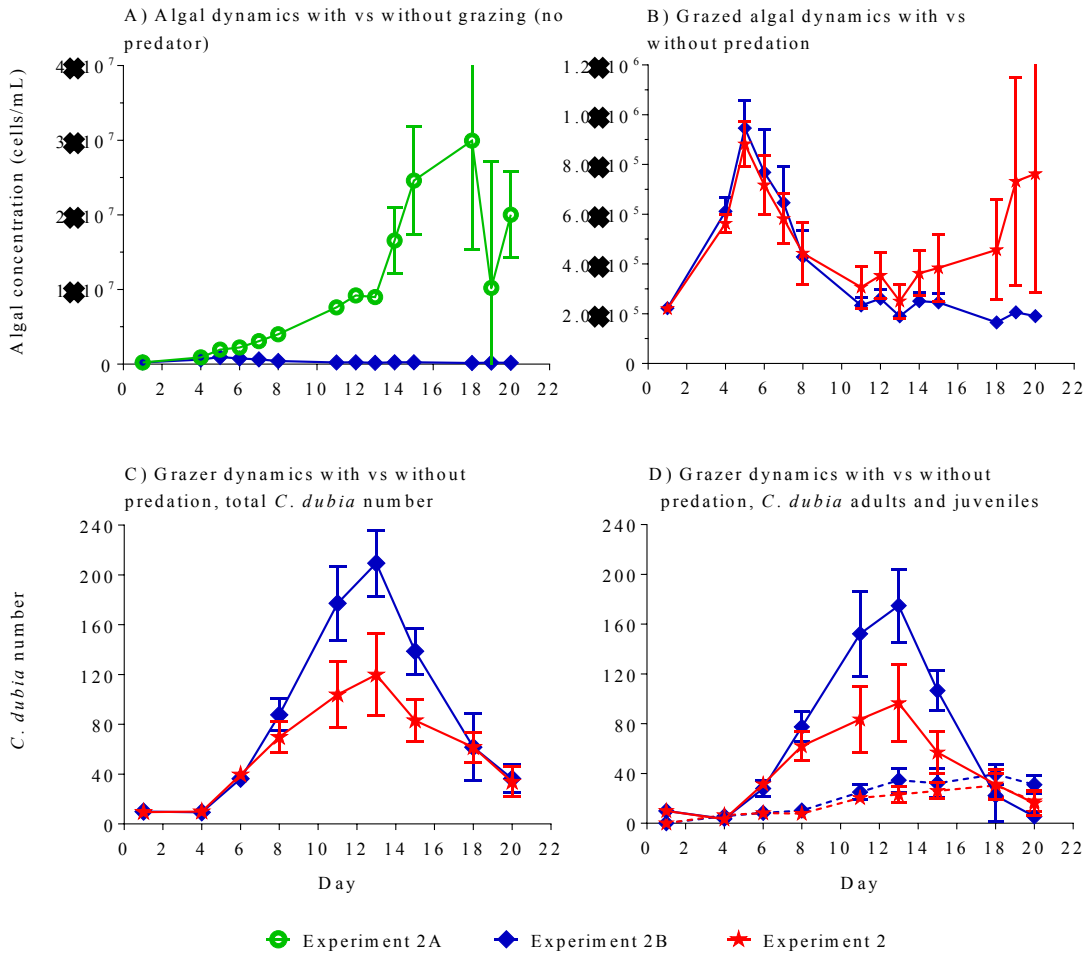

A) Algal dynamics with vs without grazing (no predator)
B) Grazed algal dynamics with vs without predation
C) Grazer dynamics with vs without predation, total C. dubia number
D) Grazer dynamics with vs without predation, C. dubia adults and juveniles
Algal concentration (cells/mL)
C. dubia number
Day
Day
Experiment 2A
Experiment 2B
Experiment 2

Figure 4