For Official Use **ENV/JM/TG/RD(2015)2**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development **05-Mar-2015**

English - Or. English

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND
THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

ENV/JM/TG/RD(2015)2
For Official Use

Test Guidelines Programme

COMPILATION OF COMMENTS RECEIVED ON THE MEDAKA MULTIGENERATION EXTENDED ONE GENERATION TEST (MEOGRT)

27th Meeting of the Working Group of National Co-ordinators of the Test Guidelines Programme (WNT)

14-17 April 2015, OECD Headquarters, Paris, France

Leon VAN DER WAL
Email : Leon.VANDERWAL@oecd.org; Tel + (33) 1 45 24 93 10

JT03371613

English - Or. English

Comments were submitted by Germany, France, the Netherlands, the United Kingdom, the United States, ICAPO and BIAC. Additional comments were received via the public OECD website by Virginie Ducrot and ILSI/HESI.

ACTION REQUIRED *The WNT is invited to take note of the responses to comments and approve the draft Test Guideline in document ENV/JM/TG(2015)4, amended as appropriate.*

No.	Country/ affiliation	General comment	Response
1	Germany	Thank you for the revised version of the now called Medaka Extended One-Generation Reproduction Test (MEOGRT). One of our main points, the 'pseudo replication' we called it, was to our understanding not discussed and we would like to ask for some clarification on that point (maybe also by a statistician). Maybe this is due to the circumstance that it was not understood what was meant by pseudo-replication. To our understanding we start in the F0 with 6 replicates of one breeding pair each. The derived eggs for the F1 are pooled and you cannot discriminate between the 6 replicates from the F1 anymore. Then the new 6 replicates of 20 eggs (that are not independent of each other anymore) are pooled again for distributing the hatchlings into 6 new replicates of 12 hatchlings each (the division of hatchings must be biased as hatchlings that seem to be affected will not be chosen). All 6x 12 hatchlings are again pooled to create 12 new replicates of breeding pairs and the last pooling takes place when distributing another 20 eggs on 6 replicates for the F2. Four times the replicates of each treatment and the controls are pooled. To our understanding that means that these replicates cannot be regarded as independent (especially with regard to hatchling pooling and division). This should also be taken into account for statistical analysis. With regard to the point that also other species are suitable for the method we like to mention Germany is about to validate a similar protocol for a zebrafish extended one generation reproducation test (ZEOGRT) that could to our understanding be integrated in the MEOGRT when ready.	Given the importance of this issue, EPA agrees with Germany that clarification should come from a statistician familiar with these types of protocols (i.e. John Green). It is our view that the pooling and redistribution steps taken do preclude comparisons within a replicate at different lifestages as there is no direct continuity within a replicate pre and post-pooling. However, this is not to say that the pooling and redistribution is in fact an example of pseudo-replication. In addition, any issues of pseudo-replication are more than outweighed by the practical requirements of running a successful test. EPA welcomes the development of an appropriate protocol for use of other fish species like the zebrafish.

			Note from Secretariat: during the VMG-Eco 10 meeting, John Green stated that for comparisons within a generation there is no issue with the so called "pseudo replication". There would be an issue if observations across generations would be compared, but this is not the case for the draft MEOGRT TG. For the comparison of observations across generations extra validation would be necessary.
2	France	Much work has been done on the set-up of this test and we thank all the people who contributed to it. After reading VMG-eco's answers to OECD members' questions, we still have some concerns regarding this guideline. I – Fish Growth The first one is the growth (weight/length) endpoint. Within each treatment, there will be 60 fish (6 replicates of 10 fish each). But since growth rate is depending on sex, growth should be treated for male and female separately. Since, fish sex cannot be determined before genetic sex determination, the sex ratio within each replicate will be totally random and skewed most of the time. Ling and Cotter (2003) worked on statistical power of experimental designs dealing with fish growth. They concluded that to get a statistical power of 80% to detect a difference of at least 10% between fish fed different diets for 3 months, no less than 200 fish per tank and 6 tanks per treatment are needed. Consistently, in most experimental studies on fish growth, hundreds of fish are	I-Fish Growth EPA believes the reference provided does not apply to a power analysis of the MEOGRT growth endpoint. For instance, the work of Ling and Cotter is based upon wild Atlantic salmon while EPA has done power analyses on actual growth data obtained from successful MEOGRTs, and these

		often used within each treatment. Therefore, with only 5 fish at best in 6 replicates, the statistical power of MEOGRT for the growth endpoint is very weak. We strongly believe that growth (weight/length) should not be considered as an endpoint in the MEOGRT. Yet, we agree that fish weight and length should be recorded. Reference: Ling E.N., Cotter D., 2003. Statistical power in comparative aquaculture studies. Aquaculture 224, 159-168. II – XX Males According to the guideline, it is expected to get naturally at least 4-5% of XX males. More XX males may be observed, but it will be difficult to judge if they are observed because of the treatment or because of other factors, because many factors may affect the percentage of XX males in a medaka population. The low number of fish used won't allow assessing for a treatment effect (except perhaps when the effect is very strong and the sex ratio is not too much skewed). Thus, how do we consider XX males in the analysis of the different endpoints? Should we discard them? This question was also asked by our colleagues of Germany in the last round, but no answer was given by VMG-eco. III – Deformities Our comment on fish deformity as a validity criterion in the control group was rejected. Yet, we strongly believe that it should be a criterion since bad handling (e.g. oil film at the water surface of the thanks, non-adapted feeding of breeders or progeny, excessive inbreeding, bad handling of eggs…) may result in massive deformity (e.g., due to failure to inflate gas bladder, lack of Ca/P in the feed, excessive content of vitamin A, genetics…) but yet not to death. In that case, such a study would have to be considered as valid because compliant with the guideline. Yet, I think we all agree that this study would not be reliable at all. IV – Developmental rate of fish between treatments Our colleagues of Germany and ourselves have pointed out the problem of eventual differences of developmental rate of fish between treatments that is not addressed by the guideline. Whilst recognizing that it is a difficult problem to deal with for the writing of the guideline, VMG-eco should be aware that hundreds of thousand euros are at stake in each case. A contingency table would	analyses indicate the current power to detect changes in growth is adequate and similar to detecting effects in other endpoints within the MEOGRT. The recommended statistics take into account that the ratio of XX to XY individuals within replicates may not be equal. II- XX Males There are two strategies to analysing the data from sex reversed fish (either XX males or XY females). 1) Censor all data from sex reversed fish across the entire test except the prevalence of sex reversal in each replicate. 2) Leave the data from all sex reversed fish in the dataset and analyse based upon genotype. *This wording was added to ANNEX 10 (stats)* III-Deformities

		therefore be very welcomed by the interested parties. V – Number of eggs at start Our colleagues from USA indicated that if the number of eggs obtained at the beginning of the experiment is too low, their number could be reduced from 20 to 15. VMG-eco agreed. Since this concerns the start of the experiment, we are more in favour of restarting the experiment rather than continuing it on a weak basis. NC comment: Please, consider comments.	EPA agrees that the prevalence of deformities could be used as a validity criterion. EPA can not provide advice on what the specific unacceptable prevalence would be, as we do not often see these deformities. IV – Development EPA is unsure of the meaning of this comment, and therefore unable to respond at this time. V-Eggs The MEOGRT protocol specifies a minimum number of breeding pairs in F0; however EPA acknowledges that during implementation, more breeding pairs per treatment and controls could be added in F0 to produce more eggs to reach the recommended 20 per replicate. *Added the last part to ANNEX 9 under the Note for Test Week 4.*
3	Netherlands	There are still the different terms 'endocrine disrupting chemicals' and	Changed to "endocrine

		'endocrine disrupters' in paragraph 1. Harmonisation of the terms is still needed in this version of the draft TG; NC comment: agree	disrupting chemicals" in para. 1
4	United Kingdom	p 1. line 8 refers to 'validation studies' - does this refer to the US-EPA 2013 in the reference list? Please add reference. p 1, line 28, spelling 'notbe' should be 'not be' p 2, under 'TEST VALIDITY CRITERA' - seem to be missing 'successful reproduction of > 65% females in controls' - please check. p 4, paragraph between line 1 and 16 seem to repeat itself twice. Please check sentence on line 2/3 and line 7/8/9, and further on line 11 and line 15/16. Also, line 12 states 'maximum solvent concentrations should not exceed 100 µg/l or 100 mg/l..' - this seem to be inconsistent with page 5, line 28 which states 10 mg/L. Please check. p 5, line 28, the reference cited 'Wheeler et al. 2013' is not consistent with the reference list which states 'Wheeler et al. 2014' - please amend accordingly. p 6, line 35, editing, please insert space between 'ANNEX5'. and 'Feeding'. p 7, line 14, editing, please remove additional dot in 'Shinomiya et al.., 2004'. p 7, line 23, insert full stop p 8, line 27-28, last sentence is a different font to the rest of document, please check. p 8, line 43, editing, 'Annex 9in' please insert space between 'Annex 9' and 'in'. p 8, line 44, refer to reference cited 'U.S. EPA, 2014?'. Please remove question mark, Also, this reference does not correspond with any of the US EPA references cited in the reference list. Please amend accordingly. p 9, line 8, editing, please insert full stop after 'etc'. p 9, line 9, editing, please remove additional dot between '..breeding pairs.' and 'During...' p 9, line 40, editing, please insert full stop. p 10, line 28, refers to 'histology guidance' - please add reference to specific guidance document. p 10, line 29, abbreviation 'RSCABS' used, please add explanation first time abbreviation used.	Fixed grammatical errors/ made suggested editorial changes. Added citations where necessary.

		p 10, line 43, editing, 'IIn additon..' should be 'In addition..' p 11, table 1, please remove commas in last column, row 2, 3, 4 and 5. p 11, line 10, editing, following '(2 wpf), remove coma and add full stop. p 12, table 2 , refer to 'thru' - suggest change to 'through to' p 12, line 11, remove additional space between 'power' and 'to'. p 13, line 3, remove the extra full stop after '(Green et al. 2014)..' p 14, line 28, editing, add come after 'anal fin papillae for F1...' p 15, line 33-34, reference McLachlan JA 2001, does not seem to be cited in the draft report, please check, and remove. p 16, line 27-28, reference Padilla et al, 2009, does not seem to be cited in the draft report, please check and remove. p 16, line 39 reference, US EPA, 2013, does not seem to be cited in the draft report, please check and remove. p 17, line 2-3, reference US EPA, in preparation. Please check so this reference is consistent with what is in the text (2014 or in preparation?). Also, currently incomplete reference 'XXXX' p 17, line 4-5, reference Wheeler et al. 2014, is cited in the draft report as Wheeler et al., 2013, - please check and amend accordingly. p 24, Annex 4, under 'Fecundity'; line 6, refer to 15 eggs - should this be 20 eggs to be consistent with Annex 3 and page 2, validity criteria? p 24, Annex 4, under 'Fertility', line 11 refers to 75% - should this be 80% to be consistent with p 2, validity criteria? p 27, Annex 7, please check legend under Figure 1, it is currently outside the page. p 30, line 20, please add full stop. p 32, line 15, in the end of the sentence please delete the additional 'MEOGRT'	
5	United States	Most of the comments are editorial clarification based on the comments received after the first round of commenting and at the December VMG-ECO meeting.	
6	ICAPO	We are grateful for the opportunity to comment on the draft TG for the MEOGRT. This is an extremely animal intensive test and additional animal welfare considerations could be implemented which we have addressed below. **Test name** It is more appropriate to call this a two generation or multi-generation	Test Name and Length of Study: Embryos of F1 generation are maintained on test only until hatch, which

		reproductive toxicity test rather than the Medaka extended one-generation reproductive toxicity test (MEOGRT) because, as mentioned in the first paragraph, the second generation is bred according to the current study protocol. Therefore, calling this test the MEOGRT may be misleading and we suggest revising the name of the test to the "Medaka Multi-generation Reproductive Test" and update the relevant parts of the TG as appropriate. **Number of fish and replicates used in the study** This is an extremely animal intensive test and everything possible should be done to minimise the number of animals used. The current draft of this TG currently requires 6 replicates in the test concentrations and 12 in the controls. We recognise that the number of controls has been increased compared to the number used in the validation study to increase the power of the study, however, we strongly suggest that this be reviewed once more experience has been gained with the study to determine if this high number of replicates is necessary. **Length of the study** Recent retrospective studies have called into question the need for a second generation for mammalian reproductive toxicity testing (Janer et al., 2007, Reproductive Toxicology 2007;24(July (1)):97–102; Piersma et al., 2011. Reproductive Toxicology. 31:392–401) and during development of the Extended One Generation Reproductive Toxicity Study (EOGRTS) by the OECD, the majority of member countries felt that the evidence supported deletion of the second generation altogether in the current study design, which contains endpoints with increased sensitivity relative to the standard two-generation test (OECD. 2011. Test No. 443: Extended One-Generation Reproductive Toxicity Study). Therefore, although we are pleased that the current draft of the test guideline indicates that the study should not be extended past hatching of the F2 generation it is still debatable whether the F2 generation should be allowed to hatch at all and whether the study should be terminated when the F1 generation lay their eggs. **Number of test concentrations**	can refer to the 'extended' reference in the title. Additionally, the name and length of the test was chosen to harmonize with the USEPA TG for medaka. At this time, the recommendation is to keep the embryos of F1 through hatch. Note from Secretariat: i.e. the embryos originating from F1 parents. Number of fish and test concentrations: While in principal EPA agrees with this comment, however in practice, implementation of this would often result in re-running the test because of not observing either a LOEC or NOEC. Once enough data are generated, the number of replicates/test concentrations recommended could be revisited. Humane endpoints:

		The current study design requires 5 test concentrations and a control (two controls if a solvent is used) with 6 or 12 replicates, respectively resulting in this being an extremely animal intensive test. We, therefore, recommend that only 4 test concentrations be required for the test. This would provide a sufficient number of concentrations for a dose response curve and, as noted in paragraph 1 ECx values are rarely suitable for large studies of this type where increasing the number of test concentrations to allow determination of the desired ECx may be impractical. **Humane endpoints** Humane endpoints (humane killing of moribund fish) should be introduced into the test guideline. Although overt toxicity should not be observed in this TG there is a risk that some fish may reach a moribund state. If this occurs, these fish should be humanely killed using an appropriate method to minimise the time that the animal suffers. This practice should already be routine in Europe under Directive 2010/63 and, therefore, please add that fish showing signs of overt toxicity should immediately be removed from the test and euthanized.	As you indicated, the practice should already be routine under the Directive, but there is concern about trying to define overt toxicity in this guideline which would result in removal of test organisms. There was a similar discussion at the Dec. 2014 VMG-ECO meeting regarding morbidity and the acute fish TG 203. A consensus for a definition of morbidity (which would result in removing a fish) could not be reached at that meeting. Therefore, additional language was not added to this TG.
7	BIAC	BIAC appreciates this opportunity to comment on the revised draft medaka extended one generation reproduction test guideline.	
8	public Virginie Ducrot	Most of the comments raised during the last meeting of VMG Eco were covered. Here I have three types of comments: - suggestions for re-wording in some places - comments on some aspects of the guideline that were not discussed during the last meeting. **- comments on some agreements that were reached during the last VMG-eco but are not yet reflected in the revised draft: these are the most important points to consider and I thus highlighted then in bold.**	
9	public ILSI/HESI	I am writing on behalf of the HESI Animal Alternatives in Environmental Risk Assessment Technical Committee, which is tasked with ensuring the	Revised/added in places in TG

		development of a sound technical basis for alternative test methods as a means to reduce, refine, or replace standard ecotoxicity test procedures around the globe. One of the recommendations stemming from the group's work is a proposed revision of life-stage / life-interval terminology in the suite of OECD Test Guidelines to ensure harmonization under a single set of nomenclature rules and decisions. There are several instances within the MEOGRT where the terminology used to describe various fish life stages is either inconsistent or incorrect. These instances are noted below. Terminology in the main text and the Annexes should be harmonized, as the terms used in both places are not consistent.	
		Paragraph 1	
10	Germany	Last sentence: please change to 'Other small fish species may be adapted to a similar test protocol. The specific methods and observational endpoints detailed in this guideline are applicable to Japanese medaka alone.' Reasoning: The ability to determine genetic sex is very specific to medaka. That would exclude zebrafish. NC comment: Agreed.	Revised wording, however, retained the point that the ability to determine genetic sex is important for this protocol as it is constructed for the medaka.
11	United States	Line 7-9. Suggest modifying the sentence: "Additional investigations would be needed to justify the utility of extending the F2 generation beyond hatching; ~~data from the validation studies has not provided sufficiently valuable justification so far~~ at this time, there is insufficient information to provide relevant conditions or criteria for warranting the extension of the F2 generation. ~~in some cases~~" Recommendation to include language, after the above sentence, that acknowledges the possibility of future guidance regarding the F2 generation as the science in this area evolves. For example:: "This Test Guideline may be updated as new information and data are considered. For example, guidance on extending the F2 generation through	Revised wording

		reproduction may be potentially useful under certain circumstances (e.g., chemicals with high bioconcentration potential or indications of transgenerational effects in other taxa)." NC comment: Concur NC comment: Concur	
12	ICAPO	Lines 10-14: Please update the sentence as follows: "The method gives primary emphasis to potential population relevant effects (namely, adverse impacts on survival, development, growth and reproduction) for the calculation of a No Observed Effect Concentration (NOEC) or an Effect Concentration (ECx), although it should be noted that ECx approaches are rarely suitable for large studies of this type where increasing the number of test concentrations to allow for determination of the desired ECx may be impractical **which causes significant animal welfare concerns due to the large numbers of animals used**."	Added wording
13	BIAC	Line 6 – "*...relevant to ecological hazard and risk assessment of chemicals, including endocrine disrupting chemicals*" suggest change to "*...relevant to ecological hazard and risk assessment of chemicals, including suspected endocrine disrupting chemicals."*	Added wording
14	public Virginie Ducrot	Line 15. This test does not directly addresses population effect, so please delete "populations"	Deleted
		Paragraph 2	
15	Netherlands	In this paragraph the following is stated: "It should be noted that if a test substance or its metabolites are not suspected of being EDCs, it may not be necessary to measure these secondary endpoints". In order to avoid unnecessary testing, this paragraph should elaborate more on when it is considered acceptable to conduct such a test as the wording "it may not be" is open to interpretation. Some companies will perhaps conduct the test just "to be on the safe side". NC comment: this open language is important because the information available	No change

		will be rather different in the various legislative frameworks No change needed	
16	Netherlands	It should be made very clear what the major differences between this test and the MMT are, as concerns the purpose of performing the test. When reading the paragraph 2 from both draft guidances there are no clear differences in the measured biological endpoints. NC comment: agree	There should not be any major differences in the endpoints measured between the MEOGRT and MMT, except that the option of extending the F2 was removed from the MEOGRT. The title was changed from "MMT" to "MEOGRT" in large part to be harmonious with the USEPA draft TG.
17	ICAPO	Lines 27-29. Please consider updating this sentence as follows: "It should be noted that if a test substance or its metabolites are not suspected of being EDCs, it may **not be** necessary to measure these secondary endpoints **and less resource and animal intensive studies may be more appropriate (OECD 2012a)**."	Added
18	BIAC	Line 28 - a space is needed between "not" and "be"	Fixed
		Paragraph 3	
19	public Virginie Ducrot	The proposed test guideline provides limited mechanistic information. Endpoint responses including decreased fecundity, decreased secondary sex characteristics, and changes in gonadal histopathology can result from non-endocrine disrupting stressors (i.e. hepatotoxicity or confinement stress) (Aluru and Vijayan, 2009; Milla et al., 2009; Pankhurst and Van Der Kraak 2000; Foo and Lam, 1993; Haddy and Pankhurst, 1999; Clearwater and Pankhurst, 1997; Lethimonier et al., 2000; Wu et al., 2003) so a weight of evidence approach should be taken when interpreting endocrine related endpoints: this was agreed during the VMG-eco meeting but it is not stated in the guideline. I recommend adding this short sentence after line 3. "A weight of evidence approach should be taken when interpreting endocrine related endpoints".	Added

		Paragraph 4	
20	ICAPO	Please update the sentence as follows to take into consideration the animal welfare implications of the study: "The test should include an adequate number of individuals to both ensure sufficient power for the evaluation of reproduction-relevant endpoints (Annex 3) **whilst ensuring that the number of animals used is the minimum required for animal welfare reasons."** Please also include the number of animals expected to be used in this test. The numbers were included in previous drafts of this TG but they have been removed from the most recent version and it is unclear why the numbers have been removed.	The animal numbers are provided in Annex 3 which is considered sufficient. Added requested additional language regarding animal welfare
21	public Virginie Ducrot	Information about the adequate number of individuals to ensure sufficient statistical power has been removed but not replaced, so that the current sentence provide no guidance about this number. I suggest adding information about this number here.	
		Paragraph 6	
22	ICAPO	The exposure period used for the F0 generation is currently set at 3 weeks. However, the exposure period should be based on the toxicokinetics of the substance. If the rate at which the chemical accumulates in the fish is known from other studies, and steady state is reached before 3 weeks then the exposure period could be reduced. Rapidly accumulating substances may not justify exposure for 3 weeks. For animal welfare reasons, paragraph 28 and annex 9 should be adapted accordingly.	Information about how the chemical accumulates not only in the adult medaka but also the transfer to developing oocytes and sperm in the breeding adults would need to be known for medaka. It is unknown how much of this information may be available. Additionally, the 3 weeks depends also on the development time of eggs and other toxicodynamic processes.
		Paragraph 7	

23	Netherlands	L 33: typo: 'a validity criterion' or 'validity criteria' L 34: If interpreted literally, this is a strange sentence. Should it not read: 'and equal treatments within the test'? Unless this statement refers to paragraph 27 on the test conditions oxygen, temperature and pH. If this is the case, this should be mentioned explicitly to avoid confusion. NC comment: agree	Revised wording regarding water temperature (Netherlands, US, BIAC).
24	United States	Line 33-35. Suggest adding clarification language at the beginning of the sentence on line 33: Regarding water temperature, while not a validity criteria, replicates within a treatment should not be statistically different from each other, and treatments within the test should not be statistically different from each other (based on daily temperature 34 measurements, and excluding brief excursions). NC comment: Concur	
25	BIAC	Line 19- 20 -Remove "throughout the test" or "over the entire duration of the study" as they are redundant. Lines 33-35 - if this statement refers to daily temperature then temperature should be in the first line. It is a confusing statement as it is currently written.	Removed
26	public Virginie Ducrot	The mean daily fecundity of the controls (F0 and F1) should be greater than 20 eggs per pair per day. This is a pretty high value: is it realistic in the Medaka when fishes are paired? I suggest lowering this value, to increase the chances of reaching the target value. E.g. align with EPA OCSPP 890.1350 (FSTRA for FHM) where a validity criterion for fecundity of 15 eggs/female/reproductive day/replicate is presented for fathead minnows. Similarly, the survival thresholds seem too high considering that manipulating young larvae is very stressful for them. I suggest setting the survival threshold to 70% to align with OECD TG 234.	These values were based on laboratory studies conducted with medaka, and based on the results across several laboratories, these minimum fecundity and survival criteria are considered adequate. Additionally, in the OECD 210 Fish Early-life stage TG (which was revised in 2013), the

			minimum survival post-hatch is also 80% for medaka.
		Paragraph 8	
27	Netherlands	For the purpose of this guideline, "sufficient reproduction" should be quantified, otherwise it is open to interpretation. The same holds for "adequate embryo survival". NC comment: agree	Then numbers needed to fill hatching incubators and evaluation of subadult sampling (which correlate to sufficient reproduction and embryo survival are listed in Para 36 and 38. Added those here in Para 8 for references.
		Paragraph 14	
28	Netherlands	In the event that use of a solvent is unavoidable, and microbial activity (biofilming) occurs, we recommend careful recording/reporting of this per tank (at least weekly) throughout the test, to improve interpretability of results and avoid unnecessary re-testing/animal use. This text could be inserted, e.g. in line 9 after 'dilution water only'. NC comment: agree	Added
29	United States	Line 7-9. The first part of this sentence may give the impression that not using a solvent may be okay if there is sufficient historical data. Further in the guideline (para. 17, line 40), stronger language regarding solvent use and historical data is used that better states the intention of the guideline at this time. For this paragraph, recommend removing "In the absence of historical data" or add language similar to paragraph 17, line 40. ~~In the absence of historical data,~~ If solvent carriers are used, appropriate	Changed

		solvent controls should be evaluated in 8 addition to non-solvent (negative) controls (dilution water only). Line 14-16. This last sentence in the paragraph seems redundant with the sentence on line 10 and 11. Suggest removing entire sentence on line 14-16. NC comment: Concur NC comment: Concur	
		Paragraphs 15	
30	Netherlands	Line 21: The period is missing in "Japanese medaka alone The" NC comment: agree	changed
31	BIAC	Line 21: a period is needed at the end of the sentence after the word "alone"	
		Paragraph 15 + 19	
32	Germany	In both paragraphs the disease treatment of the culture fish is addressed. We recommend to only addressing this once. Paragraph 15 now states: '... and no disease treatments should take place immediately before the test begins.' Immediately is not very precise. To our understanding **no** disease treatment should occur during the two week acclimation phase and during the test. Disease treatment before acclimation should – if necessary - only be allowed up e.g. 4 weeks before the acclimation starts (maybe reference can be found for a safe duration). This is also in the interest of the lab that does not want to run into any problems with this test because of stressed fish in such an expensive study. We suggest to change paragraph 19 as follows: 'Fish should not receive treatment for disease in the two-week acclimation period preceding the test and during the exposure period ~~and disease treatment should be completely avoided if possible~~ . Fish with clinical ... . '	changed

		Editorial: line 21 missing full stop … , the specific methods and observational endpoints detailed in this guideline are applicable to Japanese medaka alone. NC comment: Agreed.	
		Paragraph 17	
33	BIAC	Line 37: Suggest changing the phrase "will interfere" to "may interfere" or "could interfere"	changed
		Paragraph 20	
34	BIAC	Line 10: While it is appreciated that the culture temperature has been widened, it would still be helpful from a GLP perspective to have the statement similar to other test guidelines where "approximately" is used before the temperature. Otherwise, extensive culture records may be required to be reported in GLP studies.	While sensitivity to records management is acknowledged, given the careful consideration regarding culture and test temperature in this protocol, the suggested wording was not incorporated.
		Paragraph 21	
35	ICAPO	Please add that testing for the suitability of a solvent is only required when there are no historical data on the safe use of the solvent. It mentions that the highest concentration should not exceed the water solubility, 10 mg/L or 1/10th of the 96h-LC50. Please emphasise that it is not necessary to conduct an LC50 test using fish prior to conducting this test and other means of determining the highest concentration should be used.	At this time, another critical consideration in the recommended used of a solvent control may not be the potential influence of a solvent on the safety of a pesticide (e.g., will not cause mortality), but the potential influence on the measured endpoints for which historical data is not available. Therefore, no additional reference to historical

			data was added. Added LC50 study results to list of other possible data sources for setting doses.
36	public Virginie Ducrot	Numbers of replicates are not in line with actually existing OECD TG´s 210/229/234. I understand that the presented test design has been optimized to ensure sufficient statistical power all along the test. However, the high number of replicates reduces the number of labs which can actually run the tests (limited space in the facilities, limited staff), and the statistical power also depends on the tested chemical and shape of associated dose response curve. In order to ensure broad use of the guideline, we could insert some flexibility here e.g. "the number of replicate could be reduce for practical reasons only if it can be demonstrated that this reduction does not impair the statistical power of the test".	The increased number of replicates compared to other TGs is acknowledged. Efforts were undertaken to evaluate the optimal number of replicates to ensure adequate statistical power of the test. As this is a new TG, adequate historical data is unlikely to be available to reliably reduce the number of replicates, and it is unclear what would constitute 'impairment' of the statistical power.
		Paragraph 23	
37	ICAPO	It is mentioned that in order to ensure adequate statistical power the number of controls are doubled. When solvents are used, using the current study design, both solvent and non-solvent controls are used. Therefore, there is already double the number of controls and it should not be necessary to double the number of controls in both the solvent and non-solvent controls because, the solvent and non-solvent controls can be pooled (assuming the solvent does not impact the outcome of the test, which should be expected if the solvent has been selected appropriately). This would be a feasible option once it has been demonstrated, using historical data from the laboratory, that the solvent does	In the future, this may be something to consider as datasets become available. Although, it is acknowledged that this TG is not anticipated to be routinely requested, and therefore, may take a significant amount of

		not affect the outcome of the study.	time to collect enough data. Additionally, it is known that not all regulatory authorities support pooling of controls for statistical analysis.
38	public ILSI/HESI	Line 42: 20 eggs (fish) – change to embryos	changed
		Paragraph 24	
39	public Virginie Ducrot	Minimum weight for females & males is set as females ≥ 300 mg and males ≥ 250 mg; in EPA´s Medaka Multi Generation test validation document the following values were/are presented: females ≥ 250 and males ≥ 200 mg. Should both documents be aligned?	The draft USEPA 890 TG also has 300 and 250, therefore, no changed were made.
		Paragraph 27	
40	Netherlands	See note on paragraph 7 (same issue). Please specify for which parameter(s) the statement in line 18 and 19 refer to. NC comment: agree	Changed
		Paragraph 28	
41	Netherlands	Line 23: "F0 breeding pairs are humanely killed". Perhaps refer to paragraph 33 for further guidance. NC comment: agree	added
42	ICAPO	Please see the comment above for paragraph 6 regarding the exposure time for the F0 generation.	See above response
		Paragraph 29	
43	Netherlands	It is recommended that the flake food composition and chemical analysis is reported. It is stated that "Food with an elevated level of phytoestrogens that would compromise the response of the test to estrogen agonists should be avoided." We would like to suggest that "estrogen agonists" should be changed into "endocrine active substances"	Modified

		NC comment: agree	
		Paragraph 32	
44	United States	Line 21-23. Suggest slightly modifying sentence. Also, during the course of the histopathological assessment, the gonad is evaluated and much more powerful analyses for assessing the gonad phenotype in the context of the genetic sex ~~are done~~ are conducted. NC comment: Concur	changed
		Paragraph 34	
45	Netherlands	There are two paragraph 34, one before paragraph 33 and another after it. Please revise paragraph numbers. NC comment: agree	Modified
46	United States	Suspect this should not have received a new paragraph number, but rather be bolded and italicized, like other headings. If so, this will affect all subsequent paragraph numbering. NC comment: Concur; additionally there are errors in numbering paragraphs further down the document	
47	ICAPO	There are two paragraphs numbered 34 on page 7. Please correct numbering. Furthermore, the second paragraph labelled 34 appears to be a title or truncated sentence.	
48	BIAC	Line 37: We believe that "Handling of eggs and larval fish" is actually a header and not a new paragraph.	
		Paragraph 35	
49	ICAPO	In order to allow for natural reproductive behavior, i.e. depositing of eggs, and thus to avoid the stressful procedure of removing eggs from the female fish, spawning substrates should be provided in all tanks.	EPA has not found a substrate that all females use that would eliminate the need to collect eggs from the female.
50	BIAC	Line 2: Typo - test weeks 18	changed

		Paragraph 36	
51	BIAC	Line 5: a comma is needed after (1 pair per replicate)	Added
52	public ILSI/HESI	Line 10: hatched larvae – the term eleutheroembryo should be used to describe the stage immediately following hatching prior to the larval stage	Added
		Paragraph 37	
53	ICAPO	Criteria for death of embryos need to be defined. These could be based on the fish embryo toxicity test laid down in TG 236.	Added wording to ANNEX 9
54	public ILSI/HESI	Line 15: fertilized eggs (embryos) – the term embryo should be used here	As it is already mentioned in text, did not change
55	public ILSI/HESI	Line 18: hatchlings (young larvae) – the term eleutheroembryo should be used to describe the stage immediately following hatching prior to the larval stage	added
		Paragraph 38	
56	BIAC	Line 27: prior "to"	added
57	public Virginie Ducrot	The fish a very fragile at this life-stage. The early "handling" of the fish might result in problems with respect to the criterion of >80% survival of the larvae 3 weeks after hatch. A transfer of fish at a later stage, or a decrease in the survival threshold validity criterion would solve this problem. Since we do not go for a second generation, maybe survival threshold validity criterion in the control can be brought back to 70%, as done in OECD TG 234. **During the VMG-eco meeting, US colleagues mentioned that using a glass pipette is the safest way to handle the small fishes. Maybe this information "using a glass pipette" could be added at the end of the first sentence in this paragraph.**	This information is added in a sentence later on in para. 38. Did not add again.
58	public ILSI/HESI	Line 21 – 18, throughout: The term hatchling is used; this should be better defined as the leutheroembryo stage (see above)	It was added on para 36 and 37 for context in the rest of TG, so it was not added throughout all places where it is applicable.
59	public ILSI/HESI	Line 24: hatched larvae - the term eleutheroembryo should be used to describe the stage immediately following hatching prior to the larval stage	
		Paragraph 39	
60	ICAPO	Lines 34-35. Anesthesia should be mandatory when marking fish. Therefore, please change "Fish could be anesthetized with approved methods" to "Fish **must** be anesthetized…"	The second sentence in paragraph says that all fish are anesthetized.

		As there is evidence that fish such as medakas feel pain when having their fins clipped[1] the test guideline should stipulate that anesthesia is required prior to fin clipping. Furthermore, less painful alternatives to fin clipping are available and should be recommended and used whenever possible. For example, fish can be marked with fluorescent dyes and be recognized individually by the pattern of marks.[2] An option that may be appropriate for this test is to mark fish with dye when anesthetized. The genetic sex of the fish can be determined by taking blood samples when the fish are anaesthetized e.g. from the caudal vein if fish are not too small. [1] Sneddon (2009): Pain reception in fish: indicators and endpoints. ILAR journal 50(4): 338-342; Roques et al. (2010): Tailfin clipping, a painful procedure: Studies on Nile tilapia and common carp. Physiology & Behavior 501: 533-540. [2] Frederick (1997): Evaluation of fluorescent elastomer injection as a method for marking small fish. Bull Mar Sci, 61(2): 399-408).	Revised wording to indicate use of approved methods.
61	public Virginie Ducrot	9-10 wpf means in general sex determination after 56 dph - a FSDT is running 60 dph. At this development point the Medaka is still a very small fish. A fin clip from the caudal fin might result in insufficient DNA-material per sample. I think it could be worth to add information on the minimal quantities of tissues to collect to be able to run the PCR in this paragraph, and also to add an annex describing the method for running a PCR with a small tissue sample size. Other Point of concerns regard how the fin clipping and animal isolation will influence the reproductive behavior. Indeed, separate a swarm fish in such a small tanks might impact reproduction behavior: it could end up in lower reproduction success (and maybe changes in the endocrine system?). I guess this point of concern was already dealt with by the US colleagues, but it was not discussed during the last VMG-eco meeting. It would be good to have a feed back in this issue, focusing on the problem a how the separation of this swarm fish might impact its reproduction.	This method is described in USEPA draft 890 TG. Added reference in OECD TG. For the second comment, the fish are only held separately for a short period (about two days) until the genetic sex of the fish is established (lasting effects on reproductive behaviour is not anticipated), after which fish are selected for establishment of breeding pairs (para. 40).

		Paragraph 40	
62	Netherlands	Line 44 –Typo "(U.S. EPA, 2014?)" NC comment: agree	Citation was updated
63	United States	Line 44. The U.S.EPA 2014? Citation will need to be updated. This will apply to other places in guideline where USEPA 2014 is referenced (i.e., Paragraph 43). NC comment: Concur	
64	BIAC	Line 44: Please be sure to update the EPA reference and remove the ?	
		Paragraph 41	
65	Netherlands	Line 9 –Typo "breeding pairs. . During" NC comment: agree	Changed
66	BIAC	Line 9: There is an extra period from the removal of a sentence.	
67	public Virginie Ducrot	Why use single breeding pairs? According to other guidelines breeding groups are used. "OECD 229: Four vessels or replicates per treatment are used for Medaka (each vessel containing 3 males and 3 females)": using several breeding pairs might be a better option to preserve the reproductive behavior in a swarm fish, which is of interest when investigating endocrine disrupting effects.	The US and Japan both have conducted tests that compare paired to group spawning. Paired spawning data is superior.
		Paragraph 42	
68	BIAC	Line 18: After Test Weeks 21-26/27 should there be F2.	Removed as the test is terminated at week 19 (no continuation of F2)
69	public ILSI/HESI	Line 18: juvenile / subadult – should check this terminology against the wpf definitions in Table 1 to ensure consistency	
		Paragraph 43	
70	BIAC	Line 27: typo - remove the parenthesis Line 31: Update the EPA document reference	Changed
		Paragraph 44	
71	public Virginie Ducrot	**This paragraph does not clearly define which is the endpoint of relevance that has to be analyzed based on the data on secondary sex characteristics. This question is not solved later on: Table 1 mentions "external sex ratio" as the**	EPA recommends leaving paragraph 44 as is

		relevant endpoint. Yet, the discussions we had during the last VMG-eco meeting led to the agreement that sex reversion is the relevant endpoint for the risk assessment and should be reported. There was also a need to better define sex reversion (e.g. does it include only fish with 100% reversion or also fish with intersex or undifferentiated fishes?) and colleagues from EPA were asked to come up with recommendations how to handle that. **Therefore, I suggest:** **- adding a sentence at the end of paragraph 44 stating that "data on secondary sex characteristics should be considered together with data on the genetic sex of fish in order to assess the frequency of sex reversion." And then to provider the information obtained rfrom EPA on this question.** **- replacing "external sex ratio" by "sex reversal ratio" in table 1.**	
		Paragraph 45	
72	Germany	There are two paragraphs numbered 45. NC comment: Please correct.	Corrected
		Paragraph 46	
73	Germany	Please delete the last sentence, the test is not extended into the F2 anymore. ~~The same 21 sampling timepoints (at the completion of the exposure in Test Week 32) and processes are followed if extending the test into the F2 generation.~~ NC comment: Agreed.	Removed
74	United States	Line 18-19. "If needed, a tissue sample can be taken to repeat the *dmy* analysis to verify genetic sex of specific fish." Suggest clarifying what the additional tissue sample is comprised of (e.g., caudal fin). NC comment: Concur	The tissue type is irrelevant to the DMY analysis therefore no changes made to TG.
75	public Virginie Ducrot	If an appropriate permeabilization step is performed prior to fixation, the body cavity does not need to be opened. This should be considered as an option here	This is in draft TG
		Paragraph 47	
76	United	Line 33-34. "If needed, a tissue sample can be taken to repeat the *dmy* analysis	removed

	States	to verify genetic sex of specific fish." This sentence is a repeat of above (p 46, line 18-19). Suggest removing it here in paragraph 47 if it referring to the same tissue as in paragraph 46. If not, then suggest clarifying what the additional tissue sample is comprised of. NC comment: Concur	
77	BIAC	Line 26: Update the EPA document reference	
78	public Virginie Ducrot	**It was agreed during the VMG-eco meeting that the histology of liver and kidney should be optional endpoints, and that we should provide clear guidance on when to look at these tissues. BIAC wrote a paragraph to reflect this decision. This text has not been used in the revised version. Paragraph 47 should be modified to account for this decision of the VMG-eco.**	The language was added to paragraph 33 in the draft provided. However, additional language was added to Para. 47 to re-emphasize this.
		Paragraph 48	
79	Germany	We still propose to insert the time of spawning start as a further reproductive endpoint in the TG. The effort to record this endpoint is minimal (as eggs must counted the beginning of spawning can be easily recorded) and it provides additional and useful information on sexual development. The response that VMGeco agreed not to include that endpoint is not understandable. What are the reasons behind this decision (as the summary record of the meeting is still not available)? NC comment: Agreed.	Time to 1st spawn was listed in Table 1 for the subadults and was discussed in para 38 and ANNEX 10 (stats). However, additional wording was added in para. 32 In the experience of EPA, Spawning in controls typically starts before the fish are put into pairs. EPA/ORD records the date that spawning is observed within each replicate prior to the setup of breeding pairs. Therefore, it would not be likely linked to specific

			males and females but rather on a more general treatment group effect.
80	Netherlands	Line 43 –Typo "IIn addition" NC comment: agree	Changed
81	United States	Line 43. ~~I~~In addition… (remove extra "I" at beginning of sentence) NC comment: Concur	
82	BIAC	Line 43: Typo - IIn addition Table 1: Formatting is off in the subadult (9 or 10 wpf) - Secondary sex (there is a large space between the two words) and there should be a line above External sex ratio	
83	public Virginie Ducrot	In the first sentence "The MEOGRT provides data that can be used " please add "in a weight of evidence approach" as the endpoint under discussion are not specific to the mentioned AOPs.	added
84	public ILSI/HESI	Line 44: juveniles in 4 wpf – juvenile is defined in table 1 as 5 wpf. Need to be consistent	Changed juvenile to fish, also removed test week information as it appeared to be incorrect in places
85	public ILSI/HESI	Page 11, Table 1: The various lifestages are defined as follows: • Embryo (2 wpf) • Juvenile (5 wpf) • Subadult (9 or 10 wpf) • Adult (12 – 15 wpf) The terms "eleutheroembryo" and "larvae" should be defined in wpf for medaka in this table.	See previous response about inclusion of these terms
86	public ILSI/HESI	Page 18: All of the life-stage terms (embryo, eleutheroembryo, larvae, subadult, adult) should be included here along with the number of weeks post fertilization used for medaka.	
		Paragraph 49	

87	United States	Table 2. Suggest fixing row heights/word placement to capture words that are slightly cut-off. NC comment: Concur	Fixed table
88	BIAC	Table 2 should contain something about the F2 generation. Week 5 should have F1 and not F0 in the box and week 19 should have F2 and not F1 in the box.	
89	public ILSI/HESI	Page 12, Table 2: The following terms are used in Table 2 but a timeline (e.g., wpf) for each stage is not provided: • Embryo • Larvae • Juvenile • Subadult • Adult In addition, the table is confusing in relation to life stages, as the number of study weeks does not correspond to wpf. This table needs to be revised.	The numbers appear to be correct in the table.
		Paragraph 50	
90	public Virginie Ducrot	Is genotyping of all fish really necessary? Could we skip the F0, as these fish origin from the unexposed cultures? This would reduce stress (animal welfare) and workload. There might be some XX specimens In the F0 generation the maximum interference are possible. The investigation of the parental fish after use would result in a percentage distribution of xx fish at a later stage where there are sacrificed anyhow.	It is the opinion that genotyping all fish, including all F0 fish, is required. XX males in F0 can threaten the validity of the entire test which will only be discovered after completion or near completion of the test.
		Paragraph 53	
91	Netherlands	Line 3 –Typo "(Green et al. 2014).. " NC comment: agree	changed
92	United States	Line 1-2. Suggest replacing the first sentence in para. 53 with the following wording: ~~For fecundity, one-way ANOVA on the transformed replicate means should be~~	changed

		~~calculated from the daily egg counts, followed by Dunnett contrasts~~ For fecundity, egg counts taken daily, but may be analyzed as total egg counts or as a repeated measure. Annex 9 provides the details of how this endpoint is analyzed.	
		Paragraph 55	
93	BIAC	Line 14 – Clarify what is meant by "censored from the analysis"?	Remove the data from the statistical analysis
		Paragraph 56	
94	United States	Lines 21-26. Suggest modifying the following sentences to allow for greater flexibility in which control is used for statistical comparison to allow for differences between regulatory authorities. If statistically significant differences are detected in these endpoints between the dilution water control and solvent control groups, best professional judgment should be used to determine if the validity of the test is compromised. If the two controls differ, the treatments exposed to the chemical should be compared to the solvent control unless it is known that comparison to the dilution water control is preferred, ~~the treatments exposed to the test chemical should be compared to the solvent control, or best professional judgment should be used to determine if the validity of the test is compromised~~. If there is no statistically significant difference between the two control groups it is recommended that the treatments exposed to the test chemical are compared with the pooled (solvent and dilution-water control groups), unless it is known that comparison to either the dilution-water or solvent control group only is preferred. <u>NC comment:</u> Concur	changed
		Paragraph 57	
95	Netherlands	Line 29 –Typo “for F0, F1 .; ” <u>NC comment: agree</u>	Changed
96	BIAC	Line 19 under Test Conditions - Ammonia is not an analyte required by the draft TG (cf paragraph 27).	removed

		References	
97	United States	Add the following reference that is cited in the revised fecundity data section in the annex 10 (statistical). Cameron, A.C., and K.T. Pravin 2013. *Regression Analysis of Count Data*. Cambridge. NC comment: Concur	Added
		Annex 3	
98	Germany	Line 6: We commented on the loading rate, that it requires a specification with regard to fish density, e.g. 1 g/L. The response was 'stocking density should be supplied by the lead country' – Please supply the stocking density. Line 13. We still do not agree with the given number of fish. When calculating the minimum number of fish we get: **F0: 84 fish** **F1: 840 fish** calculation (20 eggs=embryos per replicate have to be considered) 240 fish in control (12x20) 600 fish in treatments (5x6x20) **F2: 840** (see above) Compared to 504 fish for the F1 and no fish for the F2 in Annex 3. To our understanding the 20 eggs per replicate in the F1 and F2 that are brought to hatch should be counted as fish. F1 and F2 are counted. Setup with solvent control **F0: 108 fish** 24 fish (dilution water control) 24 fish (solvent control) 60 fish in treatments **F1: 1080 fish** 240 fish in control 240 fish in solvent control	If the minimum tank size and flow rates are used as specified, then ASTM guidelines are easily satisfied. The count in the MEOGRT is correct when the unit counted is the post-eleutheroembryo.

		600 fish in treatments **F2: 1080 fish** (see above) Please rewrite as follows: 13. Number of organisms per test: Minimum of 84 fish in F0 and ~~504~~ **840** in each of F1 **and F0** (If solvent control is used, then 108 fish in F0 and ~~648~~ **1080** fish are used in each of the F1 **and F2** generations). NC comment: All points agreed.	
99	BIAC	Point 9: The inclusion of F1: continuously from fertilization is unnecessary. Age at initiation is typically the start of the test. Age at initiation needs to reflect the new text which would be >12 wpf. The limit of 16 wpf is only recommended not currently required.	Added wording from earlier in guideline about >12 but recommend to be <16
		Annex 4	
100	BIAC	Line 7 (growth): Typo - 9 weeks wpf	Removed weeks
		Annex 7	
101	United States	Figure 1: The caption for Figure 1 is skewed to the right margin (all wording still seems to be on page). NC comment: Concur	Fixed
102	BIAC	Figure 1. They have different colours for F0 and F1, but under the hatch on the right hand side should there be a new colouring for F2 which includes the boxes of 20 eggs on the right hand side of this Figure. Also, the Figure 1 legend goes off the page and it should be amended to remove the reference to the F2 generation being maintained through reproduction	Added F2 and some color
		Annex 9	
103	Germany	Our comment from the last commenting round: 'F1 generation: (line 35) Again directly after creating the mating pairs the remaining fish should be killed. See comments for the F0 generation. Why the whole test should have to be stopped and started again in case of a minor problem?' was not addressed. We suggest inserting a similar sentence at the end of the paragraph on Test Weeks 18 that was introduced for the F0 generation: 'If necessary F1 breeding pairs maybe kept for an additional 1-2 days in order to restart F2.	added

		NC comment: Agreed.	
104	United States	Line 32-34: Since the option of extending the F2 into reproduction has been removed. Suggest removing last sentence under "Test Weeks 19-20 (F2)" ~~If the exposure experiment is extended to the F2 32 reproduction phase, then the hatchings are pooled and distributed to the replicates to continue their exposure in 33 the same way as the Test Weeks 5-6~~. NC comment: Concur	removed
105	public Virginie Ducrot	*Test Week 18 (repeat of Test Week 4)* (F1 and F2) The test now stops after hatching of F2, so please delete F2 here.	
106	public Virginie Ducrot	*Test Weeks 19-20 (F2)* Please add here add a sentence about the possibility to keep the animals one or two days longer in case it is needed to restart a F2 (as done above for F0 with the sentence "If necessary F0 breeding pairs maybe kept for an additional 1-2 days in order to restart F1")	added
		Annex 10	
107	United States	Additional wording regarding additional statistical methods for analysing fecundity data that was suggested at the December 2014 VMG-ECO meeting does not appear to be in this revised draft version. It is noted that this additional language may need to be revised since the current draft does not include the option of extending F2 through fecundity, and therefore analysing with all generations included may not be applicable at this point. Additionally, the current wording in the draft guideline is the statistical methodology EPA prefers and is in the current draft OCSPP 890 guideline. Therefore, suggest replacing the paragraph labelled Fecundity Data with the following paragraphs: Fecundity Data Unless generations are to be compared, separate analyses are done for each generation, consisting of a step-down Jonckheere-Terpstra or Williams' test to determine treatment effects, provided the data are	Added with revisions removing references to generation comparisons

		consistent with a monotone concentration-response. With a step-down test, all comparisons are done at the 0.05 significance level and no adjustment for the number of comparisons made. The data are expected to be consistent with a monotone concentration response, but this can be verified either by visual inspection of the data or by constructing linear and quadratic contrasts of treatment means after a rank-order transform of the data. Unless the quadratic contrast is significant and the linear contrast is not significant, the trend test is done. Otherwise, Dunnett's test is used to determine treatment effects if the data are normally distributed with homogeneous variances. If those requirements are not met, then Dunn's test with a Bonferonni-Holm adjustment is used. All indicated tests are done independently of any overall F- or Kruskal-Wallis test. Further details are provided in OECD 2006. Alternative methods can be used, such as a generalized linear model with Poisson errors for egg counts (with no transform), if justified statistically (Cameron and Trividi 2013). Statistical advice is recommended if an alternative approach is used. Daily Egg Counts with all generations included The ANOVA model is given by Y=Generation+Time+Generation*Time+Treatment + Generation*Treatment + Time*Treatment + Generation*Time*Treatment, with random effects of Replicate(Generation*Treatment), and Time*Replicate(Generation*Treatment), allowing for unequal variance components of both types across generations. Here Time refers to the frequency of egg counts (e.g., Day or Week). This is a repeated measures analysis, with the correlations between observations on the same replicates accounting for the repeated measures nature of the data. Main effects of treatment are tested using the Dunnett (or Dunnett-Hsu) test, which adjusts for the number of comparisons. Adjustments for the main effect of generation or time are needed, for with these two factors, there is no "control" level and every pair of levels is a comparison of	

possible interest. For these two main effects, if the F-test for the main effect is significant at the 0.05 level, then the pairwise comparisons across levels of that factor can then be tested at the 0.05 level without further adjustment.

The model includes two- and three-factor interactions, so that a main effect for, say, generation, may not be significant even though generation has a significant impact on results. Thus, if a two- or three-factor interaction involving generation is significant at the 0.05 level, then one can accept the comparisons of levels of generation at the 0.05 significance level without further adjustment. Similar statements can be made for Time.

Next are F-tests for significance of treatment within time and generations, the so-called slices in the ANOVA table. If, for example, the slice for treatment within generation F1 and time 12, is significant at the 0.05 level, then the pairwise comparisons for treatment within generation F1 and time 12 can be accepted at the 0.05 level without further adjustment. Similar statements apply to tests for time within a generation and treatment and for generation within a time and treatment.

Finally, for comparisons not falling under any of the above categories, comparisons should be adjusted using the Bonferroni-Holm adjustment to p-values. Further information on analyses of such models can be found in Hocking (1985) and Hochberg and Tamhane (1986).

Daily Egg Count within a Single Generation

If separate analyses are done for each generation, a simplified procedure can be defined as described above with all references to generation removed.

Alternatively, the raw data are recorded and presented in the study report as the fecundity (number of eggs) per replicate for each day. The replicate mean of the raw data should be calculated then a square root transformation applied. A one-way ANOVA on the transformed replicate means should be calculated followed by Dunnett contrasts. It may also be helpful to visually inspect the

		fecundity data of each treatment and/or replicate with a scatterplot that displays the data through time. This will allow an informal assessment of potential effects through time. NC comment: Concur	
108	BIAC	Line 15: Typo - remove MEOGRT at the end of the sentence	removed