Emerging investigator series: use of behavioural endpoints in the regulation of chemicals

Marlene Ågerstrand, Kathryn Arnold, Sigal Balshine, Tomas Brodin, Bryan W. Brooks, Gerd Maack, Erin S. McCallum, Greg Pyle, Minna Saaristo and Alex T. Ford

Interest in behavioural ecotoxicology is growing, partly due to technological and computational advances in recording behaviours but also because of improvements of detection capacity facilitating reporting effects at environmentally relevant concentrations. The peer-reviewed literature now contains studies investigating the effects of chemicals, including pesticides and pharmaceuticals, on migration, dispersal, aggression, sociability, reproduction, feeding and anti-predator behaviours in vertebrates and invertebrates. To understand how behavioural studies could be used in regulatory decision-making we: (1) assessed the legal obstacles to using behavioural endpoints in EU chemicals regulation; (2) analysed the known cases of use of behavioural endpoints in EU chemicals regulation; and (3) provided examples of behavioural endpoints of relevance for population level effects. We conclude that the only legal obstacle to the use of behavioural endpoints in EU chemicals regulation is whether an endpoint is considered to be relevant at the population level or not. We also conclude that ecotoxicity studies investigating behavioural endpoints are occasionally used in the EU chemicals regulation, and underscore that behavioural endpoints can be relevant at the population level. To improve the current use of behavioural studies in regulatory decision-making contribution from all relevant stakeholders is required. We have the following recommendations: (1) researchers should conduct robust, well-designed and transparent studies that emphasize the relevance of the study for regulation of chemicals; (2) editors and scientific journals should promote detailed, reliable and clearly reported studies; (3) regulatory agencies and the chemical industry need to embrace new behavioural endpoints of relevance at the population level.

Environmental significance

The peer-reviewed literature contains ecotoxicity studies investigating behavioural effects such as migration, dispersal, aggression, grouping, reproduction and feeding in vertebrates and invertebrates. However, little is known about the studies’ contribution to regulatory decision-making. This study concludes that it is possible to use behavioural endpoints in EU chemicals regulation if the endpoint is assessed as relevant at the population level, but that there are few examples of such use. Recommendations for researchers, regulators, risk assessors and scientific journals who strive to improve the use of behavioural endpoints in environmental risk assessment of chemicals are provided.

Introduction

In recent years there has been growing interest surrounding behavioural ecotoxicology. For example, there has been a steady increase in behavioural effect measurements in the US EPA ECOTOX database, currently adding up to 17 324 measurements for the aquatic environment and 13 809 for the terrestrial environment. An area of particular rapid expansion is the use of fish models for drug discovery and design. This interest in behavioural endpoints is partly driven by technological and computational advances in recording behaviours, but also due to an increasing number of laboratory studies recording behavioural effects at environmentally relevant concentrations. What was in the past quite a laborious process of
careful observations and potentially watching hours of video footage is now high-throughput computer recorded endpoints without human subjectivity. Such advances are important because behavioural ecotoxicology has been previously criticized for concerns over observational errors. Technological advances, for example in electronic tagging, have also allowed the monitoring of behaviour in situ during field studies, which is helping scientists and regulators gain confidence in laboratory derived endpoints. In fact, changes in behaviour are increasingly considered valuable for advancing the next generation of ecotoxicology research.\(^7\)

Traditionally the regulatory evaluation of chemicals, including pesticides and pharmaceuticals, has not included behavioural endpoints, even though such endpoints have several advantages. These advantages include that behaviour: (1) provides a connection from molecular and physiological processes to population level processes; (2) is a sensitive ‘early warning signal’ of chemical contamination since behavioural responses can occur at lower levels of contamination than more traditional endpoints; (3) improves the ecological relevance of environmental risk assessments due to the well-established theoretical framework and suite of fitness-related endpoints underpinning behavioural studies; and (4) provides a high throughput identification of potential underlying chemical mode of actions.\(^8\)\(^9\) In addition, it can be argued that from a resource and animal welfare perspective, behavioural endpoints of sufficient reliability (i.e. inherent quality and repeatability) and relevance (i.e. appropriateness for a particular hazard identification\(^10\)), should be used in chemicals regulation, particularly considering that public health and environment research is routinely funded by public funds.

Hazard and risk assessments to decide on management measures for chemicals are common for national and international regulation agencies. Ecotoxicity studies are used to determine the potentially environmentally hazardous properties of a chemical and to support establishment of acceptable exposure levels in environmental matrices. Traditionally, regulatory assessments have been based on ecotoxicity studies measuring mortality, growth, reproduction, and development. These individual-level endpoints have been described as having a clear connection to the persistence of a population.\(^16\) To facilitate the assessment process and promote use of studies across jurisdictions, standard studies investigating these population relevant endpoints have been developed by international and national organizations, such as OECD and US Environmental Protection Agency.\(^17\)\(^18\) Inclusion of other endpoints, such as behaviour, are often seen in higher tier assessments, and/or in what is called a “weight of evidence” assessment. Ecotoxicity studies are evaluated for their reliability and relevance when deciding which studies to include in the regulatory assessment.\(^19\) Following standard procedures in ecotoxicity studies has led to reliable results applicable for chemicals regulation due to robust test protocols and detailed reporting requirements. In addition, partly due to the use of endpoints considered to be important on a population level, standard studies have been assessed as relevant for chemicals regulation. Ecotoxicity studies that are not performed according to a standard procedure and include alternative endpoints, such as behaviour, run the risk of being disregarded or seen as evidence of lower weight. Often, standard studies are performed by, or on behalf of, the regulated party, while non-standard studies are performed by researchers working in academia.\(^20\)

### Aim and methodology

Given the abundance of information available in the peer-reviewed literature on the impact of chemicals on the behaviour of non-target organisms, this study aimed to:

1. Assess the legal obstacles to using behavioural endpoints in EU chemicals regulation.

Regulatory guidance documents that set the scope for assessment of chemicals were analysed with the purpose to understand if, how and where ecotoxicity studies investigating behavioural endpoints could be used in the assessment of chemicals. Guidance documents for the following six EU policy and regulatory areas were analysed: the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) regulation; Classification, Labelling and Packaging of chemical substances and mixtures (i.e. the CLP regulation); the Biocidal Products Regulation; the Plant Protection Products regulation; Environmental assessment for human and veterinary products; and Derivation of Environmental Quality Standards (within the Water Framework Directive). The following aspects were investigated:

- If the guidance document prohibits use of behavioural endpoints.
- If the guidance document recommends or require tests using behavioural endpoints.
- What weight behavioural endpoints are given in chemical assessments.
- How population level effect is defined.
- Other aspects relevant for use of behavioural endpoints in chemicals regulation.

2. Analyse the known cases of behavioural endpoints used in EU chemicals regulation.

Due to lack of transparency in EU chemicals regulation and lack of searchable regulatory databases, it is not possible to, in an accessible way, get an overview of current use of behavioural endpoints. Instead we chose to analyse the six known, to us, cases in EU chemicals regulation. These six cases were known to us because of public discussions or due to our personal discussions with regulators and risk assessors at research institutes and national regulatory agencies. The regulatory use was examined to clarify the role of ecotoxicity studies investigating behavioural endpoints for decision-making in these specific cases. The following aspects were investigated:

- Type of behaviour endpoint used.
- Which regulatory framework it was used in.
- The weight the study was given in the chemical assessment by the risk assessors (as key, supportive or low) and how this rank was justified. Key studies are studies used when setting guideline values such as an environmental quality standard (EQS), supportive studies are considered important evidence but are not considered to be key study (e.g. due to issues related
to reliability and relevance of the study, or other studies showing toxicity at lower concentrations), and studies with low weight are the ones that are discussed but do not contribute to the overall conclusion of the assessment.

A comparison with the use in the US, and a comparison with human health assessment in EU, was made.

(3) Provide examples of behavioural endpoints with relevance for population level effects.

From a regulatory perspective, contaminant exposure only matters if it influences survival, development and reproduction because these measures can cause population decline. However, there are a number of fitness related behaviours that result in population change. Examples from both invertebrates and vertebrates were provided for different behavioural categories where there are clear links to population dynamics. These behavioural categories include migration, dispersal, aggression, sociability, reproduction, feeding, and anti-predator behaviours, and were selected because of their strong link to population growth and health. We describe how the behavioural effects of chemical contaminants, which are observed at the individual level, can have important cascading impacts and ultimately detrimental effects on population level processes.

Results
Regulatory guidance on the use of behavioural studies in EU chemical assessments

None of the analysed EU guidance documents prohibited use of behavioural endpoints when assessing environmental effects (Table 1). In four of the policy areas, behaviour endpoints were not mentioned as examples of endpoints of interest. Instead, there was a general focus on traditional endpoints such as survival, reproduction, and development. Two of the guidance documents for the REACH regulation stated that behavioural studies could be used as supportive evidence, but not without backing from studies using more traditional endpoints. Examples of such behavioural endpoints included sediment avoidance or burrowing activity.

The guidance document for deriving EQS states that behavioural endpoints are unsuitable as the basis for EQS derivation since the endpoints “do not include direct measurements of survival, development or reproduction”. However, the guidance document makes an exception for changes in behaviour resulting in impaired competitive fitness and avoidance reactions that would make individuals avoid contaminated habitats where otherwise they normally would be present. In contrast to the guidance document for deriving EQS, the guidance documents for Plant Protection Products regarded behavioural endpoints as key evidence. Examples of this are the assessment of field data for bees, and acknowledgement of avoidance and behaviour related to reproduction as important endpoints when assessing risks for birds and mammals.

For the four policy areas where the regulated party is obliged to perform new ecotoxicity studies if existing studies are not sufficient to fulfill the data demands, it is recommended to use standard studies, primarily from OECD. Currently, observed changes in behaviour need to be reported according to several OECD standard studies. For example, the OECD TG 222 Earthworm Reproduction Test states that an inability to dig into the soil should be reported. In the OECD TG 246 Acute Contact Toxicity Test for bumblebees, signs of reduced coordination should be reported. However, very few standard studies are designed to directly investigate effects on specific behaviours; the behavioural observations made are instead an additional measure for the primary endpoints studied, i.e. effects on mortality, growth, reproduction and development. The clearest examples of situations where behavioural endpoints are considered important in OECD standard studies, together with other endpoints, are when assessing effects on bees using OECD TG 213, 214 or 245. In contrary, there are two standard studies from ASTM International specifically addressing behaviour: E1604 for testing aquatic organisms, and E1768 for testing freshwater fish. For human health assessment, the OECD TG 426 Developmental Neurotoxicity Study on rats is used to investigate behavioural endpoints such as motor activity, learning and memory, anxiety, and social behaviour.

Several of the EU regulatory documents that we analysed mentioned that there is a need to develop new standard studies with behavioural endpoints. For Plant Protection Products, standard studies for birds that include endpoints such as avoidance behaviour, parental care and nesting behaviour, are requested. For bees, a standard study investigating chronic toxicity for larvae is missing, the guidance document recommends using an extended version of OECD TG 213 where behavioural endpoints are included. Currently, the OECD is developing the standard study “Homing flight test on honeybee after single exposure to sublethal doses”.

A principle within regulatory assessments of chemicals has been to base decisions on toxicity and ecotoxicity studies investigating adverse effects. WHO defines adverse effects as “Change in the morphology, physiology, growth, development, reproduction or lifespan of an organism, system or (sub)population that results in an impairment of functional capacity to compensate for additional stress or an increase in susceptibility to other influences”. What separates human health assessments from environmental assessments is that population relevant endpoints and not individual effects are considered in the latter. The only known exception to this level of a protection goal in EU is a recommendation from EFSA to protect aquatic vertebrates (fish and amphibians) at the individual level in acute risk assessments to avoid visible mortality.

In the outdated and no longer used Guidance Document on Risk Assessment for Birds and Mammals, the focus on population relevant effects was interpreted as direct effects on survival, development or reproduction, and therefore guidance like the following can be found: “One aim of the ecological risk assessment is to predict effects on the population level, although this is difficult or impossible to measure directly. The usual approach is based on the consideration that effects on populations will not occur if the survival rate, reproduction rate and development of individuals are not affected. Therefore, in principle, only endpoints in toxicity tests which are related to these key factors of population dynamics are ecotoxicologically relevant”. However, in more recently developed guidance documents populations have
<table>
<thead>
<tr>
<th>Responsible agency and policy/regulatory area</th>
<th>Guidance document(s)</th>
<th>Recommendations from the guidance documents</th>
<th>Example text from the guidance document (the number represent guidance document in column 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>European Chemicals Agency. REACH regulation</td>
<td>(1) Part B hazard assessment (2) Chapter R.2 information requirements (3) Chapter R.4 evaluation of available information (4) Chapter R.7b + R.7c Endpoint specific guidance</td>
<td>Behavioural endpoints mentioned in two documents, where behavioural studies can be used as supportive evidence. Standard studies are recommended when performing new studies. All available studies are recommended for evaluation</td>
<td>(3) “Behavioural endpoints like sediment avoidance or burrowing activity have not been standardised. Such endpoints can give indications on toxic effects but should not be interpreted in isolation” (4) “Reproduction tests include parental and reproductive endpoints. An endpoint relating to overall reproductive success should normally be selected to define the long-term NOEC. Depending on the individual case and the availability of data, this could be the reproduction rate, the survival or growth rate of the offspring, or behavioural parameters in adults or young”. “Screening endpoints such as behavioural responses, i.e. avoidance testing should not be interpreted in isolation”</td>
</tr>
<tr>
<td>European Chemicals Agency. Classiﬁcation, Labelling and Packaging (CLP) regulation</td>
<td>Guidance on the application of the CLP criteria (1) Volume IV: environment. Part A: information requirements (2) Volume IV: environment – assessment and evaluation (parts B + C)</td>
<td>Behavioural endpoints not mentioned. All available studies are recommended for evaluation</td>
<td>—</td>
</tr>
<tr>
<td>European Chemicals Agency. Biocidal Products Regulation</td>
<td>Guidance document on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters (1) Guidance document on terrestrial ecotoxicology, SANCO 2002 (2) Guidance document on risk assessment for birds and mammals, EFSA journal 2009 (3) Guidance document on risk assessment for birds and mammals, EFSA journal 2009 (4) Draft EFSA guidance document on the risk assessment of plant protection products on bees</td>
<td>Behavioural endpoints mentioned for bees, birds and mammals. Standard studies are recommended when performing new studies. All available studies are recommended for evaluation</td>
<td>(2) “Key parameters which may be considered in a field trial include mortality (assessed via the use of dead bee traps), behaviour (including foraging behaviour in the crop and around the hive), honey crop (assessed via weighing the hive at appropriate intervals) and state of colony (including an assessment of brood)” (3) “Granivorous birds and mammals may be able to distinguish treated seeds from non-treated seeds and may show a preference for either treated or untreated seeds in their diet. This may be inﬂuenced by various factors including appearance, taste or surface texture of the treated seed, and aversive reactions to the active ingredients”</td>
</tr>
<tr>
<td>European Food Safety Authority, the European Parliament and Council. Plant Protection Products Regulation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responsible agency and policy/regulatory area</td>
<td>Guidance document(s)</td>
<td>Recommendations from the guidance documents</td>
<td>Example text from the guidance document (the number represent guidance document in column 2)</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>----------------------</td>
<td>---------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>European Medicines Agency. Medicinal Products Regulations</td>
<td>(1) Guideline on the environmental risk assessment of medicinal products for human use (2) Guideline on environmental impact assessment (EIAS) for veterinary medicinal products – phase I and II (3) Guideline on environmental impact assessment for veterinary medicinal products in support of the VICH guidelines GL6 and GL38</td>
<td>Behavioural endpoints not mentioned. Standard studies are recommended when performing new studies. All available studies are recommended for evaluation</td>
<td>“...the assessor may be faced with data from studies describing endpoints that do not include direct measurements of survival, development or reproduction but, rather, describe e.g. behavioural effects, anatomical differences between control and treatment groups, effects at the tissue or sub-cellular level, such as changes in enzyme induction or gene expression. Generally these are unsuitable as the basis for EQS derivation. However, some other endpoints are relevant. For example, anatomical changes to gonad development that would prevent successful reproduction, or changes in behaviour if the effect described would impair competitive fitness may be relevant. Avoidance reactions may also be relevant if populations are likely to avoid a contaminated habitat where they would normally be present”</td>
</tr>
<tr>
<td>European Commission. Deriving Environmental Quality Standards</td>
<td>Technical Guidance for Deriving Environmental Quality Standards. Guidance document no. 27</td>
<td>Behavioural endpoints mentioned, some can be used. All available studies are recommended for evaluation</td>
<td>“...the assessor may be faced with data from studies describing endpoints that do not include direct measurements of survival, development or reproduction but, rather, describe e.g. behavioural effects, anatomical differences between control and treatment groups, effects at the tissue or sub-cellular level, such as changes in enzyme induction or gene expression. Generally these are unsuitable as the basis for EQS derivation. However, some other endpoints are relevant. For example, anatomical changes to gonad development that would prevent successful reproduction, or changes in behaviour if the effect described would impair competitive fitness may be relevant. Avoidance reactions may also be relevant if populations are likely to avoid a contaminated habitat where they would normally be present”</td>
</tr>
</tbody>
</table>
been linked to abundance/biomass and individuals have been linked to behaviour/survival/growth.\textsuperscript{24}

There are two guidance documents explicitly stating that behavioural endpoints are relevant to the protection goal. The EFSA’s guidance document on the risk assessment of plant protection products on bees list behaviour as key factor, and it does so already in the regulation (EC 1107/2009).\textsuperscript{25} The new EFSA guidance for the identification of endocrine disruptors stresses that behavioural endpoints are implicitly covered by the WHO definition of adversity since these type of effects will have implications on reproduction and development.\textsuperscript{26}

In comparison to the EU, the Canadian Council of Ministers of Environment and some provincial jurisdictions, such as the British Columbia Ministry of Environment and Climate Change Strategy (BCENV) have established water quality guideline-development protocols that allow for behavioural endpoints if ecological relevance can be established. Such endpoints include predator avoidance, swimming ability, or certain olfactory-mediated behaviours that can be linked to ecological relevance. Ecological relevance, then, is defined as follows: “Ecological relevance pertains to whether physical abilities (e.g., swimming speed, orientation ability, and migratory fitness), physical traits (e.g., fin size/shape), physiological traits (e.g., production of a certain enzyme), and/or behavioural tendencies (e.g., swimming in groups) of organisms are important enough to influence a species’ ecological competitiveness. Characteistics that are of high ecological relevance are those that have a strong positive or negative influence on survival, reproductive ability, and growth (e.g., stunting, high fertility, and organ failure).”\textsuperscript{27} In addition, BCENV continues to monitor the literature for contaminant-induced olfactory impairment (including sensitive olfactory-mediated behaviours) to ensure that existing water quality guidelines remain protective against these sublethal effects (A. Azizishirazi, \textit{Pers. Comm.}, 2019). Thus, there are protocols and precedents that could be exported to other jurisdictions in order to facilitate the integration of behavioural endpoints into both established and newly developed regulatory frameworks for chemicals.

**Known cases of behavioural endpoints used in EU chemicals regulation**

The total number of cases where behavioural endpoint has been used in EU chemicals regulation is not known, and due to lack of transparency and searchable databases in the EU chemicals regulation it is not possible to make a quantitative analysis in an assessable way. Based on public and personal discussions with regulators and risk assessors we conclude that the overall regulatory use of behaviour endpoints seems to be rare. The six cases presented below are the only known publicly available cases to us (Table 2). From the analysis below we could conclude that there is spatial, temporal and product-specific variation in how these behavioural studies have been used in chemical registration and pre-market assessments in the different EU regulations.

**Assessed and used as key study.** A study investigating avoidance behaviour in \textit{eel}\textsuperscript{29} was used in the European Union Risk Assessment Report (RAR) from 2002 for the solvent and fuel component methyl tertiary-butyl ether (MTBE).\textsuperscript{28,29} This assessment was performed under the Existing Substances Regulation, the forerunner of the REACH regulation. The avoidance study was performed since, according to the RAR, there was a need for further information and/or testing. The

**Table 2  Examples of current use of behavioural endpoints in environmental assessments of chemicals in EU**

<table>
<thead>
<tr>
<th>Substance &amp; regulation</th>
<th>Reference</th>
<th>Behaviour endpoint &amp; test animal</th>
<th>Importance for the regulatory decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methyl tertiary-butyl ether (MTBE) in existing substances regulation</td>
<td>28</td>
<td>Avoidance (attraction) in fish (\textit{A. anguilla})</td>
<td>Key study</td>
</tr>
<tr>
<td>Methyl tertiary-butyl ether (MTBE) in Danish national EQS-derivation</td>
<td>28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-Nonylphenol in REACH (substance evaluation, SVHC)</td>
<td>32</td>
<td>Spawning in fish (\textit{P. promelas})</td>
<td>Supportive evidence but not key endpoints</td>
</tr>
<tr>
<td></td>
<td>33</td>
<td>Feeding, social and aggression in fish (\textit{O. mykiss})</td>
<td></td>
</tr>
<tr>
<td></td>
<td>34</td>
<td>Sexual in fish (\textit{P. reticulata})</td>
<td></td>
</tr>
<tr>
<td>Lambda-cyhalothrin in the Plant Protection Product Regulation</td>
<td>36</td>
<td>Drifting in insect species (\textit{B. rhodani} &amp; \textit{L. cf. Fusca})</td>
<td>Supportive evidence but not key endpoints</td>
</tr>
<tr>
<td></td>
<td>37</td>
<td>Pre-copulatory behaviour in amphipod (\textit{G. pulex})</td>
<td></td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>Drifting in invertebrates (\textit{L. nigra}, \textit{G. pulex}, \textit{H. sulphurea})</td>
<td></td>
</tr>
<tr>
<td>Alpha-cypermethrin Field study from chemical company</td>
<td></td>
<td>Foraging activity in honeybee (\textit{A. mellifera})</td>
<td>Low weight, other endpoints showing limited or low effects were considered more important</td>
</tr>
<tr>
<td>Di-2-ethylhexyl phthalate (DEHP) in the Existing Substances Regulation Zinc pyrithione in the Biocidal Products Directive</td>
<td></td>
<td>Predation efficiency in dragon fly larvae (\textit{Aeshna})</td>
<td>Low weight, due to low reliability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Avoidance in amphipod (\textit{M. affinis})</td>
<td>Low weight, due to insufficient reporting of results. However, used to justify selection of assessment factor</td>
</tr>
</tbody>
</table>

same study was, in 2009, used by the Danish Environmental Protection Agency to derive an EQS under the Water Framework Directive. In both cases, the behavioural study was considered to be the key study.

**Assessed and used as supporting evidence.** Three non-standard studies investigating different behavioural endpoints in fish[33–34] were included in the assessment report from the European Chemicals Agency (ECHA) for the industrial chemical 4-nonylphenol. A so-called “substance evaluation” was performed under the REACH-regulation with the purpose to investigate if 4-nonylphenol was to be considered a “Substance of Very High Concern” (SVHC) due to its endocrine disrupting properties. The behavioural studies acted as supportive evidence but were not considered key studies in the assessment.

In the assessment for the insecticide lambda-cyhalothrin in the Plant Protection Product Regulation three non-standard studies investigating behaviour endpoints were used.[35–36] The behavioural studies were considered to have relevant results but due to a number of factors the studies were only used as supporting evidence. Factors mentioned in the assessment were: the tested formulation was not identical to the assessed formulation; effects were observed at all treatments and hence no NOEC value could be obtained; unclear reporting of methodology; concerns regarding the study reliability; the study was a non-standard study.[37]

**Assessed as low weight due to low effect.** The Swedish Chemical Agency assessed a product containing the insecticide alpha-cypermethrin. A short-term reduction in foraging activity of dragonfly larvae exposed to the plasticizer di-2-ethylhexyl phthalate (DEHP)[38] under the Existing Substances Regulation. However, since studies investigating mortality only showed a limited and short-term increase, and since no effects occurred on the brood nest or bee brood development, the behavioural effects were not considered sufficient to alter the conclusion from the other studies. The behavioural endpoints were given low weight in the assessment due to the limited effects shown in the study.

**Assessed as low weight due to low reliability.** The Swedish Chemicals Agency used a study that measured the predation effects of dragonfly larvae exposed to the plasticizer di-2-ethylhexyl phthalate (DEHP) under the Existing Substances Regulation. However, the study was considered to be of low reliability due to the use of only one test concentration and the high amounts of ethanol used when spiking the sediment, and therefore the study was given low weight in the risk assessment.

**Assessed as “not assignable” due to insufficient reporting.** A study investigating avoidance to zinc pyrithione in a sediment living amphipod was evaluated by the Swedish Chemicals Agency in an assessment of copper pyrithione (zinc and copper pyrithione are considered to be the same substances at low concentrations, dominated by the free ion of pyrithione). This was done under the Biocidal Products Directive (BPD), the forerunner of the Biocidal Products Regulation (BPR).[39] The study was assessed as “not assignable”, using the reliability evaluation categories first developed by Klimisch et al.,[40] due to insufficient reporting of results. The study only presented an EC50-value and the raw data needed to calculate a NOEC could not be located by the author of the study. In the risk assessment document, it was also mentioned that the tested concentrations had not been analytically confirmed in the study. Despite this, the study was used as supporting evidence to justify the choice of assessment factor when deriving the Predicted No Effect Concentration (PNEC): “The assessment factor 100 seems to be most applicable, given that there is only one test (the Hyalella) where concentrations were measured over the whole exposure period, and given that there are only two organisms types tested (for mortality), and finally, also considering that the non-guidance research test with avoidance (Monoporeia) as endpoint indicate effects are expected at ~100 times lower”. To clarify, in this assessment the behavioural study was both disregarded and used as justification. This raises concerns regarding the consistency of the assessment methodology.

**Use of behavioural endpoints in the US regulation.** An example from the US shows use of behavioural endpoints in another type of chemical assessment. The US Environmental Protection Agency included behavioural endpoints in pilot “Biological Evaluations” for the data rich pesticides chlorpyrifos, malathion, and diazinon. These assessments were performed to analyse impacts on species listed under the Endangered Species Act (ESA). As introduced above, behavioural endpoints for fish, amphibians, invertebrates, birds, and mammals are described as appropriate evidence for individual level and population level, e.g. from the assessment for diazinon: “This endpoint is considered relevant to the fitness of an individual because limited locomotion would potentially increase the likelihood that an individual would be susceptible to predation as well as an inability to fly and thus migrate”. Still, behavioural endpoints are included as one line of evidence in a weight of evidence approach, and thereby acting as supporting evidence in these ESA assessments.[41]

**Use in human health risk assessments.** Use of behavioural endpoints (in animals) in human health risk assessments is more common, but still, studies investigating behavioural effects are not part of the core dataset for neither pesticides nor biocides. Instead, they are requested if triggered by results from studies on lower tiers in the assessment structure.[42,43] The use of behavioural endpoints in human health risk assessments are primarily facilitated through two OECD standard studies, the two-generation reproduction toxicity study (OECD TG 416) and the developmental neurotoxicity study (OECD TG 426). In addition to other endpoints, these standard studies measure reproductive behaviour and neuro-behavioural endpoints, respectively. Their regulatory use has not been straight forward. For example, in the assessments for the insecticides deltamethrin and lambda-cyhalothrin the use of OECD TG 426, or a similar standard study, raised concerns regarding the sensitivity of the strain of rat (another strain of rat had shown to be more sensitive in other studies) and the exposure level in offspring was not clear (exposure through lactation). An additional assessment factor was suggested to account for the uncertainties, but this was ultimately not accepted. In the end, the studies were used as supporting evidence in both assessments.[44,45]

The difficulties interpreting the results from OECD TG 426 was acknowledged in a study aimed at identifying areas of...
improvement in this standard study. The OECD TG 426 offers several possible options in the test design when assessing learning and memory. For example, active avoidance of an unpleasant action can be tested, or the ability to find a way through a maze. These two test designs assess different types of behaviour. Experienced evaluators and thorough guidance are needed to be able to detect study designs that likely will result in negative results. The flexibility in the test design was considered the main reason for the interpretation difficulties, however, the flexibility was also considered necessary due to the large variation in tested substances. An analysis of neurotoxicity studies for bisphenol A showed that effects were more often observed in the behavioural endpoints not required (e.g. social and sexual behaviour) according to OECD TG 426, compared to the endpoints required (e.g. motor activity). This was seen especially at low doses, and exemplifies the importance of non-standard behavioural endpoints in regulation of chemicals.

Non-standard studies investigating behavioural endpoints have also contributed to regulatory decisions. When the restriction proposal for bisphenol A in thermal paper (used for cash register receipts) was adopted by the European Commission, effects on behaviour (alteration of spatial memory and learning functions) was one of four identified key risks for the unborn child, together with effects on the female reproductive system, effects on the mammary gland, and risk for obesity. However, since the available studies, including the behavioural studies, did not allow for a quantification of the dose–response relationship, the studies could not be used directly for the derived no-effect level (DNEL, i.e. the level of exposure to a substance above which humans should not be exposed). Instead, a study investigating kidney effects in mice was used as key study together with the standard assessment factor recommended in the regulation and an additional assessment factor (of six) to take into consideration the effects seen in the behavioural studies.

Relevance of behavioural endpoints for population level effects

The behaviour of an organism is intrinsically linked to its ‘fitness’ in a variety of ways through its capacity to communicate, find food, evade predation or catch prey, find mates, defend territories, or undergo large scale migrations. Ecologists have been linking behaviour with population level effects for decades but within the field of ecotoxicology the need for incorporating ecologically important behavioural measures has only been articulated more frequently in recent years. Behavioural ecologists have studied how behavioural variation among animals leads to differential fitness, measured as effects on life-history characters (e.g. growth rate, development rate, survival). For example, reduced mobility (i.e. activity) has repeatedly been linked with reduced growth and development of a wide range of organisms. Another behaviour where variation has been tightly linked to fitness is sociality, an individual’s tendency to associate spatially with conspecifics. More social individuals generally experience increased likelihood of survival in the presence of lethal predators. These are just two types of behaviours (out of several) where trait-variation has been tightly linked to changes in life-history characters and as such translates into fitness (i.e. individual-level) and population effects. However, identifying which specific behaviours within or among various species are concretely linked to changes at population level due to chemical exposure has not received extensive attention. Below are examples where behavioural effects on individual level in ecotoxicology have been linked with effects and responses of population level importance.

Migration and dispersal. Although changes in animal activity rates are one of the most commonly measured behavioural responses to chemical contamination in the laboratory setting, very few studies have connected these measures of small scale activity to dispersal or migration in the wild. Recently, Hellström et al. found that the anxiolytic pharmaceutical, oxazepam, caused Atlantic salmon (Salmo salar) smolts to migrate faster both in laboratory migration pools and downstream in a river. This anxiolytic-induced boldness, however, was not beneficial in the wild because exposed salmon were predated upon to a greater extent. Similarly, Woodman and colleagues found that virile crayfish were strongly attracted to sertraline, and that sertraline-exposed crayfish showed increased aggression towards control animals. This result also suggests a maladaptive response that could increase the risk of predation in the wild. Marentette et al. found that invasive round goby (Neogobius melanostomus) collected from environments heavily contaminated with metals, polychlorinated biphenyls (PCBs), and polyaromatic hydrocarbons (PAHs) dispersed slower through a maze in the laboratory compared to gobies collected from a relatively cleaner site. However, by using mark-recapture, the authors also found that fish from the same contaminated and clean sites showed similar dispersal rates in the field, underscoring the importance of also measuring behaviour in the wild when possible. Animals migrate or disperse for many reasons, including to avoid unfavourable environmental conditions, to access more productive habitats, or to reproduce. Migration and dispersal success are therefore important behaviours that can shape the structure and viability of populations.

Aggression, sociability and reproductive behaviours. Chemical contaminants are known to impact complex social interactions such as sociability (association of two or more conspecifics in space), aggression, and reproductive behaviour, and these behaviours have direct implications for fitness and population dynamics. For example, killifish (Fundulus diaphanus) exposed to environmentally relevant doses of 4-nonylphenol form looser shoals, and this was linked to disrupted chemical signals emitted by the signalling fish. The antidepressant fluoxetine has been shown to disrupt the integration of pheromone cues to control sexual behaviours in male goldfish (Carassius auratus). Further, contaminants that impact conspecific detection can also change aggression and dominance hierarchies among individuals. Sloman and colleagues found that rainbow trout (Oncorhynchus mykiss) exposed to cadmium, which damages the olfactory epithelium, were less aggressive towards an unexposed rival and this reduction in aggression allowed groups of 10 fish to form stable
Predator-prey interactions have important implications for population-level responses to contamination. Nontargets can also affect reproductive behaviours as a result of exposure to endocrine disrupting compounds has been documented in a range of species. For example, exposure to organochlorine pesticides reduced herring gull (Larus smithsonianus) and merlin (Falco columbarius) offspring incubation time, nest attendance, and defence behaviours as a result of the complex interactions of these pesticides with acetylcholine and estrogen signalling.

In the wild, disruption of reproductive behaviours as a result of exposure to endocrine disrupting compounds has been documented in a range of species. For example, exposure to organochlorine pesticides reduced herring gull (Larus smithsonianus) and merlin (Falco columbarius) offspring incubation time, nest attendance, and defence behaviours as a result of the complex interactions of these pesticides with acetylcholine and estrogen signalling.

Feeding and anti-predator behaviours. Chemical contaminants can also affect interactions between organisms and therefore impact foraging for prey and anti-predator responses. Predator-prey interactions have important implications for animal fitness and provide a link between individual-level and population-level responses to contamination. Sources of contamination in the wild, such as wastewater effluents, are often complex mixtures of compounds that may have many mechanisms of action beyond the endocrine system. As such, researchers have also documented reduced aggression and increased courtship behaviours following exposure to wastewater.

Automated response to stimuli (high-throughput behavioural analysis). An increasing number of studies are making use of automated high throughput devices that stimulate experimental subjects to evoke behaviours such as startle responses. These startle responses can be used a proxy to determine the efficiency of predatory escape response, however, more work needs to be done to link these ‘in a box’ experiments with survival responses in wild. There are an increasing number of platforms on the market which allow for high-throughput analysis of behaviours using video cameras (e.g. ToxTrac, DaniovisionTM, ZebrawaxTM and ZantiksTM) whilst others make use of an organism’s movement through electrical conductivity. Some platforms can stimulate organisms through lighting, vibration/noise and electricity and can measure an enormous amount of different behaviours from basic speed/distance to turn angles and rates of acceleration. Longer term experiments can be conducted so that more complex behaviours can be studied such as learning and memory. Others can be adapted to multiple ways to measure attraction or repulsion from various stimuli. Within the fields of clock gene
biodiversity much has been made of devices that utilize beam-breakers (TriKinetics) of infrared light in apparatus which can be adapted to the size of the organisms. Other labs are generating lab-on-a-chip technologies which can also be adapted in various ways for the endpoints of interest. These technologies have the capacity to increase the throughput of behavioural research with not only increasing speed, productivity and ultimately lower research costs, but also by reducing the subjectivity of behavioural research.

**Discussion**

Chemicals regulation is based on a combination of scientific knowledge, pragmatic considerations and policy decisions. With advancements in science and changes in community values comes new regulations, albeit with the delay that can be expected in a democratic society where legal certainty is important.

Some might argue that the precautionary principle, a cornerstone in the EU chemicals regulation, could be used to justify the use of behavioural endpoints in environmental risk assessments. According to the European Commission, the precautionary principle should be used when there is evidence of potential harm, but there is not enough information to determine the risk with sufficient certainty. However, according to a recent review of the operation of the REACH regulation, risk management actions evoking the precautionary principle have been limited. This is because scientific uncertainties are not assessed to the extent needed for decisions based on the precautionary principle.

This study concludes that the only legal obstacle to the use of behavioural endpoints in EU chemicals regulation is whether an endpoint could be considered to be relevant at the population level or not. Since there is no agreed method for establishing this connection, this decision must currently be taken case-by-case. Here, the individual assessor and/or the responsible institution can play a large role. A broader interpretation of what is relevant to populations may result in more behavioural studies being considered for chemicals regulation.

A consequence of allowing for additional studies is that more resources are needed for literature search, evaluation and assessment. Since the majority of studies investigating behavioural effects are non-standard studies with possibly new test designs, methodologies and test species, increased complexity in the study evaluations is expected. Absence of standardized methods likely prevents assessors/institutions from including behavioural studies in chemical assessments.

This study also concludes that the current use of ecotoxicity studies investigating behavioural endpoints in regulation of chemicals is low. We do not claim to provide a complete overview of the current use of behavioural endpoints but since many regulators and risk assessors we discussed the matter with had difficulties remembering if behavioural studies had been used in chemicals regulation, stating that they are rarely used is appropriate. The available examples show that behavioural studies have been evaluated for use in several different regulatory frameworks, and that studies have been used as both key and supporting evidence. The available examples also show that studies have been disregarded due to low reliability as well as insufficient reporting. This is a problem that applies to peer-reviewed studies in general, and not behavioural studies in particular, and it can be improved by scientific journals introducing reporting requirements and by increased awareness of the regulatory system among academic researchers.

What cannot be captured by this study are the studies that were disregarded in the early steps of the assessment process, e.g. because they were not searched for, or they were disregarded because the endpoint was not considered relevant.

Interestingly, behavioural studies are used in the regulation of pesticides as evidence of the efficiency of the substance. For example, behavioural endpoints are used when assessing repellents for ticks, mosquitoes, and birds. This appears important because pesticides are designed to be biologically active compounds that elicit effects through specific mechanisms of action. For plant production products, the most common use of animal behaviour studies is as evidence that the natural behaviour of animals will prevent them from being exposed to chemicals. For example, studies showing that certain birds and rodents prefer to avoid short grass due to lack of food and hiding places are used by chemical companies as evidence that the animals will not be exposed to pesticides used on greens at golf courses. This use of behavioural studies shows that these endpoints are accepted as evidence in pesticide assessments, both by regulators and by the chemical industry.

Finally, this study concludes that there is evidence that behavioural endpoints are relevant for population level effects.

This conclusion is supported by a recent theme issue of the Philosophical Transactions of the Royal Society B where the connection between animal behaviour and dynamics of populations and communities was highlighted. The editorial stated that "Behavioural ecology has accumulated a rich toolbox for quantifying how the main behaviours of animals relating to foraging, predation, mating, parental care, communication and sociality are affected by the current threats to biodiversity, notably habitat loss and fragmentation, overexploitation, climate change, pollution, disease, and invasive species. This provides a firm foundation for a bottom-up approach to understanding human impacts on the natural world."

In summary, there are three principal obstacles to improved use of behavioural endpoints in environmental risk assessments of chemicals. Below these obstacles are discussed together with possible ways forward (Fig. 1).

**1 Lack of promotion of behavioural endpoints in chemicals legislation**

In EU chemicals regulation, the chemical industry is obliged or recommended by legislators to submit particular types of ecotoxicological studies. The regulatory guidance documents mention either specific standard studies, which currently lack behavioural endpoints (with a few exceptions), or specific effects, focusing on survival rate, reproduction rate and developmental effects. Consequently, the chemical industry is not encouraged to submit studies investigating behavioural endpoints, even though such endpoints might be equally or more relevant for a particular chemical. A way forward is to...
adjust the guidance regarding the recommended effects of relevance, whenever justified (see obstacle 3).

Future efforts are warranted to develop standard behavioural studies for regulatory applications, or to add behavioural endpoints to already established standard studies. To accomplish this, it will be important to catalogue historical and more contemporary laboratory behavioural responses to contaminants. From this, information on specific behaviours of representative taxa should be identified that are predictive of ecologically important adverse outcomes in the field. Research will be needed to accomplish this goal. These efforts should initially employ specifically acting chemicals that elicit, or are anticipated to cause, behaviour changes through specific pathways because such observations could provide diagnostic positive controls for future studies of other chemicals. Laboratory studies aimed at uncovering the degree and nature of variability should be conducted to ensure that experimental design considerations (e.g., statistical power), repeatability and practices are adequate for regulatory adoption.

(2) Low use of non-standard studies in chemicals regulation

Since very few standard studies measure behavioural effects, an immediate increased use of behavioural endpoints in environmental risk assessments would depend on the use of non-standard studies. To increase the use of non-standard studies legislation must encourage or demand inclusion of such approaches, when available. One example of this practice is reflected in the draft of the new guideline on the environmental risk assessment of medicinal products for human use: “Behaviour is an example of an ecotoxicological endpoint not yet established as a reliable and standardised endpoint. It may however be very relevant for neuro-active substances and when standardised guidelines become available, be taken up in a tailored risk assessment scheme for neuro-active substances”. However, it should be noted that adding non-standard studies to environmental risk assessments puts additional demand on those evaluating studies at chemicals companies or regulatory agencies due to the variety of test designs and test organisms used. Here regulatory guidance that promotes thorough and robust evaluations is needed.

Non-standard studies are primarily performed by academic researchers, during curiosity driven or exploratory research, as opposed to standard studies performed by, or on behalf of, the chemical industry to comply with test demands in chemicals regulations. Recent debated cases show that the use of non-standard studies in chemicals regulation have not been without obstacles, often resulting in disqualification of studies. Common reasons why non-standard studies have been assessed as “not reliable” are due to shortages in the test design or in information about how it was performed. For example, some studies are considered unreliable because of a lack of measured tested concentrations, too few replicates, or too few tested concentrations. But other studies have also been assessed as “not assignable” due to a lack of transparency in the description of the methodology and results. However, these problems have a workable solution that does not necessarily request that the original research idea is compromised. In fact, the scientific quality of the study can even benefit from it. Researchers can adjust the methodology, performance and presentation of the study in minor ways thereby ensuring that the regulatory requirements are fulfilled, e.g. by increasing the statistical power by adjusting the number of replicates and by improving the transparency of the reporting.

(3) Lack of clarification of the importance of behavioural endpoints at the population level

When a behavioural study is disqualified for regulatory use because the endpoint used is not considered relevant at the population level, the possibility for the individual researcher to influence the potential regulatory impact of the study is low. Theoretically, a detailed scientific explanation of how a behavioural change in an individual can result in effects at the population and/or community level would be sufficient. A

---

Fig. 1 Schematic overview of the risk management process and how the use of behavioral endpoints can improve in decision-making. Environmental legislation and guidance documents stipulate which studies that need to be included in the risk assessment and provide guidance for the study evaluation process. The environmental risk assessment consists of standard and non-standard studies of sufficient reliability and relevance. A risk management decision is taken, partly based on the conclusions from the risk assessment.
conceptual framework for this has been developed, intended for researchers that would like to expand their investigation from individual level to community level. Tools like Systematic Reviews and Systematic Maps has been suggested to facilitate use of behavioural studies in conservation management and policy. In EU and the U.S. there is a growing interest for such methodology in chemicals regulation.

Currently, to claim that behavioural endpoints are relevant at the population level (according to the current EU regulation), their connection to survival rate, reproduction rate and developmental effects needs to be established. Consequently, there may be a lower regulatory threshold for behavioural endpoints closely related to these effects, e.g. as shown in the use of behavioural endpoints in two standard toxicity studies used in human health risk assessments to assess reproductive and developmental effects (OECD TG 416 and 426). The results in the present study show that what is considered relevant at the population level may change between jurisdictions. For example, some legislation has opened up for use of some behavioural endpoints while others only recommend the more traditional endpoints. What is considered population relevant may also change over time. As one example, adverse histopathological effects on the kidney was recently used as key evidence when an EQS for the pharmaceutical diclofenac was set. Harmonisation among regulatory frameworks represents a useful goal.

Conclusions and recommendations

There are three main conclusions from this study:

1. Behavioural endpoints representing a wide range of behaviours are relevant at the population level and should therefore be used in chemicals regulation.

2. There is currently a low use of ecotoxicity studies investigating behavioural endpoints in chemicals regulation.

3. There are no apparent legal obstacles to be able to use behavioural studies in EU chemicals regulation, except that the endpoints need to be considered relevant at the population level.

In order to improve the current use of behavioural studies in regulation of chemicals, all relevant stakeholders are recommended to contribute:

- Researchers need to write robust and reliable papers, and also carefully explain the relevance of the study for regulation of chemicals.
- Editors and scientific journals need to promote and demand detailed study reports and the highest level of study reliability and transparency, including responsible data sharing.
- Regulatory agencies and the chemical industry need to embrace new behavioural endpoints and develop standardized methods for behaviours when relevant at the population level.

Disclaimer

This paper is the work of the authors and the opinions herein are those of the authors alone and do not necessarily reflect the view or opinions of their institutions.

Authors contributions

Ågerstrand devised the study and drafted the manuscript. All authors contributed to the manuscript and approved the final manuscript. All data is available in the manuscript.

Conflicts of interests

The authors have no competing financial interests to declare.

Acknowledgements

Thanks to researchers, risk assessors and regulators working at the Swiss Ecotox Centre, the Netherlands National Institute for Public Health and the Environment, the UK Environment Agency, the US Environmental Protection Agency, the Danish Environmental Protection Agency, the German Environment Agency and the Swedish Chemicals Agency for valuable discussions.

References


30 European Chemicals Bureau, *Addendum to the Environmental Risk Assessment for tert-butyl methyl ether (MTBE)*.

31 Danish Environmental Protection Agency, *Determining quality standards for the aquatic environment, in Methyl tertiary-Butyl Ether (MTBE), (CAS no. 1634-04-4)*, Danish, 2009.


35 European Chemicals Agency, *Support document for identification of 4-nonylphenol, branched and linear, as substances of very high concern because due to their endocrine disrupting properties they cause probable serious effects to the environment which give rise to an equivalent level of concern to those of CMRs and PBT/vPvBs*, 2012.


Environmental Science: Processes & Impacts


99 M. Pelli and V. P. Connaughton, Chronic exposure to environmentally-relevant concentrations of fluoxetine (Prozac) decreases survival, increases abnormal behaviors, and delays predator escape responses in guppies, *Chemosphere*, 2015, **139**, 202–209.


120 M. Ågerstrand, S. Christiansen, A. Hanberg, C. Rudén, L. Andersson, S. Andersen, *et al.*, A call for action: Improve reporting of research studies to increase the scientific basis for regulatory decision-making, *J. Appl. Toxicol.*, 2018, **38**(5), 783–785.


