Introduction

Ecologists have often treated conspecific individuals as broadly equivalent, after accounting for age, sex and other group effects (Bolnick et al. 2003, Yamamoto et al. 2014, Wakefield et al. 2015). However, there is increasing realisation that individuals in many animal populations differ substantially in resource use (‘niche variation’), and the term ‘individual specialisation’ has been used to describe heterogeneity in resource use (Araújo et al. 2011, Patrick et al. 2014, Ceia & Ramos 2015). The extent of niche variation has important implications, including the potential not only to reduce the degree of intra-specific competition, but also to increase individual efficiency in finding and handling food (Estes et al. 2003, .

*Corresponding author: ana.bertoldi.carneiro@gmail.com
**These authors contributed equally to this work
Cook et al. 2006, Kotzerka et al. 2011, Ceia & Ramos 2015). Moreover, if a wide range of habitats and resources are used, a particular species or population may be better buffered against anthropogenic impacts, as individuals are likely to respond in different ways to changes in the environment (Tinker et al. 2008, Phillips et al. 2009, Dias et al. 2011, Masello et al. 2013). Hence, the characterisation of divergent behaviour or strategies may provide important insights into the ecology, evolution, conservation and management of the species (Thieman et al. 2011, Wakefield et al. 2015, Ramírez et al. 2016).

The interest in individual variation, however, extends far beyond the marine ecology literature, and has evolved independently in several fields, resulting in a fragmented literature with different terminology (Dall et al. 2012). While some researchers have focused on individual niche specialisation (particularly in terms of behavioural traits associated with foraging behaviour and diet choice), others have focused on documenting behavioural syndromes or animal personalities (boldness, aggressiveness, activity, exploration and neophobia), and the division of labour within insect societies or hunting groups (Gazda et al. 2005, Dall et al. 2012). Although these similar concepts (individual niche specialisation, behavioural syndromes or animal personalities, and the division of labour) are applied in different contexts, they are largely concerned with the same behavioural properties, and each field uses similar statistical methodology for describing individual variation (Cleasby et al. 2015).

Individual specialisation occurs when individuals use a narrow subset of the ecological niche of the population, for reasons not attributed to any a priori ways of classifying individuals (Bolnick et al. 2003, Woo et al. 2008, Votier et al. 2010, Dall et al. 2012, Patrick et al. 2014). Indeed, there are multiple biological reasons to distinguish between phenotypic variation that is sex- or age-related and individual-level specialisation (Bolnick et al. 2002). In marine predators, sex differences may arise from the influence of size dimorphism on sex-specific parental roles, intersexual competition, foraging and locomotory efficiency (including diving capability), or habitat specialisation (González-Solís et al. 2000, Shaffer et al. 2001, Phillips et al. 2004a, Breed et al. 2006, Staniland & Robinson 2008, Quillfeldt et al. 2011, Stauss et al. 2012); these mechanisms are potentially, but not necessarily, different from those generating individual specialisation. Similarly, differences in behaviour are often associated with variation in foraging abilities (i.e. competitive ability) or performance of adults of different ages or experience (Navarro et al. 2010). After accounting for effects of sex and age, other sources of variation among individuals may still exist; this residual variation is what describes individual specialisation. The unexplained within-individual variation is generally assumed to be distributed normally and uniformly between individuals when using this approach (Westneat et al. 2015).

Although a variety of approaches have been used to detect individual specialisation in traits (e.g. diet and foraging behaviour), most approaches rely on contrasting the amount of variation within individuals with the variation between individuals, using repeatability analysis. A repeatable behaviour will show relatively low within-individual variance compared to between-individual variance (Bell et al. 2009, Nakagawa & Schielzeth 2010, Dingemanse & Dochtermann 2013). Consistently divergent behaviours have also been termed specialisations. The definition of behavioural consistency, however, is not trivial when looking at the statistical methodology, and there is still no universal agreement (but see Cleasby et al. 2015). In the context of specialisation, consistency relates to the within-individual variation. The within-individual variation needs to be compared to the between-individual variation in order to test for specialisation (to determine a ‘high’ vs. a ‘low’ within-individual variance). Although behavioural consistency can reflect specialisation, it does not necessarily indicate individual specialisation as defined here. For example, if groups of individuals (e.g. males vs. females, adults vs. juveniles) differ in their preferences, but these group differences are not taken into account, then researchers might wrongly conclude that there is individual specialisation.

Individual specialisation and behavioural consistency are known to be widespread across a range of taxa and behaviours; they have been demonstrated for a number of species of marine predators in foraging behaviour, migratory routes, dive characteristics, diet, timing of events, activity patterns and habitat choice, and foraging site fidelity during breeding and non-breeding periods (Holz et al. 1989, Staniland et al. 2004, Croxall et al. 2005, Phillips et al. 2005, 2006, Sargeant et al. 2005, Guilford et al. 2011, Thiebot et al. 2011, Patrick et al. 2014, Yamamoto et al. 2014, Wakefield et al. 2015). Because marine predators forage on highly patchy and more or less predictable food resources at the coarse- to meso-scale (tens to hundreds of kilometres) and over days and weeks, the incidence of specialisation is unsurprising (Weimerskirch 2007, Ceia & Ramos 2015, Wakefield et al. 2015). The recurrent use of similar areas may
increase familiarity with feeding conditions, including fine-scale resource availability and distribution (Hamer et al. 2007, Ramírez et al. 2016). Marine predators will often target regions characterised by local physical features or processes, including eddies, frontal systems, upwelling zones and shelf breaks, that increase primary production or serve to aggregate various types of prey (Kappes et al. 2010, Louzao et al. 2011, Pinet et al. 2011, Baylis et al. 2012, Arthur et al. 2015, Wakefield et al. 2015). Foraging behaviour and diet specialisations may therefore emerge as a result of the spatial and temporal availability and predictability of prey (Woo et al. 2008, Navarro & González-Solís 2009, Patrick et al. 2015, Sommerfeld et al. 2015). Moreover, because marine predators frequently target such productive areas in the breeding and non-breeding seasons, and are central-place foragers during breeding, competition may be high and lead to the use of divergent foraging strategies (Estes et al. 2003, Villegas-Amtmann et al. 2008, Patrick et al. 2014, Ceia & Ramos 2015).

Several conventional approaches, such as visual observations and the analysis of pellet, midden, regurgitated or stomach contents, have been used to test for, or to quantify, consistency in diet in marine predators (Votier et al. 2004a,b, Hamer et al. 2007, Maldini et al. 2010, Scheel & Anderson 2012). More recently, these approaches have been replaced or supplemented by the use of biologging technology and stable isotope analysis of carbon and nitrogen (Phillips et al. 2005, Furness et al. 2006, Anderson et al. 2009, Newsome et al. 2009, Matich et al. 2011, Kernaléguen et al. 2015). The aim of this review is to scrutinise the most common methods used to detect and quantify individual specialisation and behavioural consistencies in foraging, movement patterns and diet of marine predators, particularly seabirds and marine mammals (Table 1). A series of fixed factors unrelated to individual specialisation per se are also discussed briefly, as these need to be taken into account when applying several of the approaches presented here. The review focuses on 3 types of data: those derived from (1) diet assessed using conventional approaches, (2) diet assessed using stable isotopes, and (3) tracking. In each case, we present the various statistical analyses used to date, providing an overview of the specific advantages and disadvantages.

**DIET USING CONVENTIONAL APPROACHES**

Many methods are used to study marine predator diet. Traditionally, these involve the collection and analysis of regurgitated prey items and pellets, stomach contents, faeces, direct observations of prey caught or carried by returning adults, or dropped items collected at breeding colonies (Pierce & Boyle 1991, Barrett et al. 2007, Moreno et al. 2016). The analysis of diet specialisation depends not only on the number of individuals sampled but also, and most importantly, on the number of independent feeding events recorded per individual in order to calculate an index of diet variation (Araújo et al. 2011). Monte Carlo simulations are the only way to do a power analysis, but some prior knowledge of the expected level of variability in a given population is necessary. Although repeated observations over time are required to quantify specialisation correctly, the first studies testing (and rejecting) the null hypothesis that conspecifics share an identical resource distribution were derived from cross-sectional data collected from a population at one specific point in time (e.g. Baltz & Morejohn 1977, Lonne & Gabrielsen 1992).

Unlike this snapshot approach, longitudinal studies, where sampling at intervals reflects multiple feeding events of the same individual, allow specialisation to be properly quantified at a particular temporal scale.

The simplest approach to detect specialisation is to identify groups of specialists and generalists based on the proportion of each dietary source in the total diet, and assign individuals to these groups (e.g. Pierotti & Annett 1991, Annett & Pierotti 1999, Oro et al. 2005, Hamer et al. 2007). For example, by observing prey items being delivered to pigeon guillemot Cepphus columba chicks for an average of 4 full days over multiple years, Golet et al. (2000) defined specialists as individuals whose diet contained more than 50% of a particular item or class of items. Using a higher threshold for the definition of specialist predators, Votier et al. (2004a) categorised great skuas Stercorarius skua as specialist fish or bird predators when the contents of regurgitated pellets collected over multiple periods between egg-laying and chick-fledging comprised 70% or more of the respective prey, or as generalists when none of the prey items comprised 70% or more of the diet. When data on diet are not available, behavioural information (such as dive depths and the proportion of trips associated with fishing boats) or the use of areas where particular prey occur (proportion of time at each feeding site) have been used to assign individuals to different diet groups (e.g. Ropert-Coudert et al. 2003, Montecvecchi et al. 2009, Masello et al. 2013, Granadeiro et al. 2014, Patrick et al. 2015, Tyson et al. 2015). The threshold used for separating specialists and generalists is often arbitrary. Ideally, researchers could...
deal with this issue by testing the sensitivity of their results to different cut-offs. Furthermore, although this approach can be used to infer differences in diet between individuals, which can be interpreted as a form of individual specialisation, it does not allow estimation of its statistical significance.

To solve these issues, Bolnick et al. (2002) introduced a framework to quantify and test statistically for individual specialisation in diet that has since been used in a wide range of taxa. By using ratios of variance in a continuous trait measured for each prey item (e.g. prey size), it is possible to estimate how much of the total population variance is explained by differences within vs. between individuals. In practice, the average intra-individual variance (‘within-individual component’, WIC) and the intra-individual variance (WIC + BIC = TNW, the ‘total niche width’). This index (WIC/TNW) varies from 0 (complete individual specialisation) to 1 (no individual specialisation). Empirical values can be compared with a null model (randomisation of prey items between individuals). This approach can be implemented in the R package RInSp (Zaccarelli et al. 2013). The method can also be extended to discrete data such as the frequency of alternate prey in the diet by using diversity indices as a proxy for variance (Bolnick et al. 2002). Instead of the raw number (or mass) of diet items, values are transformed into a proportion matrix. Examples of diversity indices applied for this purpose are the Shannon-Weaver index and a modified version of Hill’s ratio (Golet et al. 2000, Tinker et al. 2008, Woo et al. 2008). Alternat-

Table 1. Summary of the most common methods used to detect and quantify individual specialisation and behavioural consistencies in foraging, movement patterns and diet of marine predators. WIC: within-individual component, GLMM: generalised linear mixed-effects model, TNW: total niche width, UDOI: utilisation distribution overlap index

<table>
<thead>
<tr>
<th>Method</th>
<th>Data type</th>
<th>Example</th>
<th>Reference</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparing categories based on percentage of use</td>
<td>Categorical</td>
<td>Prey items</td>
<td>Votier et al. (2004a)</td>
<td>Does not allow statistical testing of the degree of individual consistency</td>
</tr>
<tr>
<td>Ratios of variance (repeatability)</td>
<td>Continuous variable</td>
<td>Prey trait (e.g. size)</td>
<td>Woo et al. (2008)</td>
<td>Using GLMMs allows the inclusion of fixed effects (not when using Bolnick’s WIC/TNW ratio)</td>
</tr>
<tr>
<td>Stable isotope ratios</td>
<td></td>
<td></td>
<td>Garcia-Tarrasón et al. (2015)</td>
<td></td>
</tr>
<tr>
<td>Trip summary statistics</td>
<td></td>
<td></td>
<td>Wakefield et al. (2015)</td>
<td></td>
</tr>
<tr>
<td>GLMM with random slopes</td>
<td>Response variable and continuous environmental variables</td>
<td>Use/availability locations</td>
<td>–</td>
<td>Resource selection function (habitat selection)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Foraging trip characteristics, e.g. path straightness</td>
<td>Patrick et al. (2014)</td>
<td>Random slope model</td>
</tr>
<tr>
<td>Distances</td>
<td>Spatial data</td>
<td>Between distribution centroids</td>
<td>Navarro &amp; González-Solis (2009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Between migratory routes</td>
<td>Dias et al. (2011)</td>
<td></td>
</tr>
<tr>
<td>Overlaps</td>
<td>Ranges</td>
<td>In geographical space</td>
<td>Wakefield et al. (2015)</td>
<td>Bhattacharyya’s affinity, UDOI, etc.</td>
</tr>
<tr>
<td></td>
<td>Utilisation distributions</td>
<td>In environmental space (each variable at a time)</td>
<td>Wakefield et al. (2015)</td>
<td>Bhattacharyya’s affinity, UDOI, etc.</td>
</tr>
</tbody>
</table>
Diet (trophic position/carbon source) using stable isotopes

Stable isotope analysis of carbon ($^{13}$C/$^{12}$C, $\delta^{13}$C) and nitrogen ($^{15}$N/$^{14}$N, $\delta^{15}$N) are used increasingly to test for consistency in trophic level and foraging habitat over multiple timescales. In marine ecosystems, the ratios of stable carbon and nitrogen isotopes in consumer tissues reflect those of their prey during tissue formation in a predictable manner (Phillips et al. 2009). Carbon stable isotope ratios mainly reflect the foraging habitat or carbon source of the consumer, whereas nitrogen stable isotope ratios indicate trophic position (Bearhop et al. 2000, Cherel et al. 2006, Ceia et al. 2012). Since different tissues turn over at different rates, each integrates diet information over various temporal scales (Bearhop et al. 2006, Matich et al. 2011, Ceia & Ramos 2015). For instance, plasma retains information from a few days prior to sample collection, and red blood cells from the previous 3 to 4 wk, whereas feathers and fur represent diet during moult, since keratin is metabolically inert after synthesis (Bearhop et al. 2006, Ceia et al. 2012, Barquete et al. 2013). Pinniped (otariid) and sea otter Enhydra lutris vibrissae, cheironian shells, baleen plates and mammalian teeth sampled sequentially can be used to represent several years of diet information (Hobson & Sease 1998, Walker & Macko 1999, Cherel et al. 2009, Vander Zanden et al. 2010, Eisenmann et al. 2016). Consequently, the similarity between stable isotope ratios measured in different tissues with different turnover rates or during different periods can be used as a proxy for individual diet and habitat specialisations (Wakefield et al. 2015).

$\delta^{13}$C or $\delta^{15}$N are often analysed separately; however, because $\delta^{13}$C has a trophic component, the studentised residuals of the relationship between $\delta^{13}$C and $\delta^{15}$N can be included in models testing for short- and long-term consistency to control for the variability in $\delta^{13}$C that could be due to $\delta^{15}$N (Votier et al. 2010, Ceia et al. 2012). Alternatively, a multivariate model which allows direct modelling of the covariance/correlation between these traits can be used to control for the variability in $\delta^{13}$C due to that in $\delta^{15}$N; however, this approach, to our knowledge, has not been used in the marine predator literature to date.

The most common approach, applicable to different types of data (stable isotope ratios, trip metrics, habitat use), includes the use of repeatability ($R$) analysis (i.e. intraclass correlation coefficient, ICC), where the level of individual specialisation in a population can be estimated as the proportion of the total variance accounted for by differences among individuals, according to the following formula:

$$R = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}$$

with $\sigma_a^2$ being the between-individual variance and $\sigma_e^2$ the within-individual variance. The way to estimate the variance components will depend on the distribution of the data (Nakagawa & Schielzeth 2010). For Gaussian data, such as ratios of carbon and nitrogen, repeatabilities can be calculated either with the $F$ table of an ANOVA with individual identities as fixed factors (e.g. Hamer et al. 2001, Gray et al. 2005, Vander Zanden et al. 2013, Oppel et al. 2015) or with linear mixed-effects models with individual identities treated as a random effect (e.g. Dias et al. 2011, Matich et al. 2011, Grist et al. 2005, Van der Zanden et al. 2013, Oppel et al. 2015). Note that the framework proposed by Bolnick et al. (2002), examining the within-individual component and between-individual component of a niche, is essentially the same as the residual variance (within-individual variance) and between-individual variance when using mixed-effects models. Indeed, the ratio WIC/TNW is very similar to the equation for repeatability, except that the numerator represents the within- rather than between-individual variation. Linear mixed-effect models have the advantage of directly estimating the variance necessary for the calculation of repeatability; the quantities $\sigma_a^2$ and $\sigma_e^2$ can be extracted from the output of a mixed model. For non-Gaussian data (e.g. binary, proportion and count data), generalised linear mixed-effects models (GLMMs) with the appropriate link function are required (e.g. García-Tarrasón et al. 2015, Potier et al. 2015). The advantage of using mixed-effect models to calculate repeatability is that additional covariates can be included as fixed effects to account for known sources of variation. However, controlling for fixed effects will affect variance component, and hence repeatability, estimates (Wilson 2008; adjusted repeatabilities, Nakagawa & Schielzeth 2010). The inclusion of predictors associated with individual observations (season, year, etc.) will tend to increase the repeatability, while predictors associated with the individual-level component (sex, age, etc.) will tend to decrease repeatability (Wilson 2008, Nakagawa & Schielzeth 2010). ANOVA
tables use the ratio of the sums of squares to estimate repeatabilities (Lessells & Boag 1987). The statistical significance of the repeatability can be tested with a randomisation test: measurements (e.g. isotope ratios) are reshuffled many times between individuals and repeatability is calculated for each randomisation. The corresponding p-value is the proportion of randomisations that produce a repeatability index greater or equal to the observed repeatability. This can be implemented in the R package rptR (Nakagawa & Schielzeth 2010). The importance of modelling differences in within-individual variation — e.g. using (double) hierarchical generalised linear models, or (D)HGLMs as presented in Cleasby et al. (2015) — has been recognised in a variety of fields and, recently, by ecologists, but is not discussed in detail in the present review because, as far as we are aware, it has featured only in a few studies of marine ecology (but see Cleasby et al. 2015 for a review of indices).

In theory, when data are normally distributed and the design of the study is balanced (equal number of measurements for each individual), an even simpler approach is to test for a correlation (Spearman’s or Pearson’s correlation) between pairs of measurements taken at different points in time for each individual (correlation-based repeatability; Nakagawa & Schielzeth 2010). If there are more than 2 measurements per individual, correlations can be calculated between each pair of measurements, and averaged for each individual. The main issue with correlations is that they ignore other factors, and so this might lead to within-individual consistency being high because the measurements were taken under the same conditions, rather than because the individual is consistent (Dingemanse et al. 2012).

A general question when calculating repeatability, or any other index of individual specialisation, is what are the appropriate sample sizes. Wolak et al. (2012) calculated confidence intervals for a number of repeatability estimates presented in the literature and found that for most of these studies, the associated precision was low because of inappropriate sample sizes. They provide guidelines for estimating how many individuals, and how many measurements per individual, are necessary to get a certain level of precision. In general, the higher the value of the repeatability index, the fewer the number of measurements and individuals needed for reasonable precision. Wolak et al. (2012) also emphasise that the same precision can be achieved with different combinations of number of individuals or number of measurements. In the particular case of stable isotopes, when there are often only 2 measurements per individual, the number of individuals sampled will need to be high. The formula provided by Wolak et al. (2012) for this estimation can only be applied for repeatabilities calculated using the variance components of a 1-way ANOVA table, hence for Gaussian data. For more complex model structures or non-Gaussian data, no such formula exists. However, power analyses can still be carried out at the level of each variance component in a mixed-model framework (potentially with both random intercepts and random slopes), using the R package pamm (Martin et al. 2011).

### BIOLOGGING STUDIES

A variety of devices have been used to examine animal movements, including platform terminal transmitter (PTT), GPS, and geolocator or global location sensing (GLS) loggers. These devices have different performances, and the general trade-off is between temporal resolution, deployment duration, device mass and cost (Wakefield et al. 2009). PTTs can provide multiple locations per day with accuracy typically <15 km (Burger & Shaffer 2008, Phillips et al. 2007, Costa et al. 2010). Due to their high cost, these devices have, to an extent, been replaced in the last decade by GPS loggers. The latter have a much better spatial accuracy (within 10 m) and temporal resolution (up to 1 Hz) (Guilford et al. 2008, Phillips et al. 2007, Kotzerka et al. 2010). In diving predators that only surface for short periods, very rapid (<100 ms) acquisition Fastloc GPS is required, and is slightly less accurate than conventional GPS (50% of locations within 36 m; Dujon et al. 2014). Because of the very high temporal resolution, fine-scale behavioural information can be inferred from movement (Guilford et al. 2008, Freeman et al. 2010). The use of miniaturised GPS loggers, however, is still limited by the short lifespan (weeks) of devices without solar panels. The use of GLS loggers avoids some of these problems as they have low power requirements, and are small enough to be attached long-term to a ring on the tarsus or a flipper tag (Wilson et al. 2002, Phillips et al. 2004b, Shaffer et al. 2005, Staniland et al. 2012). This technology is unsuitable for fine-scale spatial analysis, but is ideal for monitoring large-scale movements during the non-breeding season or over extended periods. Amongst the disadvantages, GLS loggers will provide only 2 locations per day with an average accuracy of 186 ± 114 km, and latitude is difficult to estimate from light for 3 to 4 wk around the equinoxes (Phillips et al. 2004b, Shaffer et al. 2005). Other devices can be deployed that collect
immersion, acceleration, temperature, images or other data allowing more detailed investigations into at-sea activity (Phalan et al. 2007, Mackley et al. 2010, 2011, Gutowsky et al. 2014, Kernaléguen et al. 2015).

Biologging studies can be used to assess the potential specialisation or flexibility of individuals from within a population because individuals can be tracked across multiple trips or over the course of a year or longer (Pinaud & Weimerskirch 2005, Soanes et al. 2013, Muller et al. 2014). As such, the data can be used to examine repeatability in foraging destinations (i.e. site fidelity), migration schedules (timing and duration of events), fidelity to wintering areas and routes, and consistency in habitat use or preference (Croxall et al. 2005, Phillips et al. 2005, 2006, Guilford et al. 2011, Thiebot et al. 2011, Baylis et al. 2015a, Yamamoto et al. 2014, Arthur et al. 2015).

Spatial analyses: distance between centroids of distributions or migratory routes

It is possible to study individual specialisation in space, and not only in trip characteristics, based on distances between the centroid of the locations at 2 different times for the same individuals tracked during the breeding season (Navarro & González-Solis 2009, Ceia et al. 2014), or between centroids in different winters (Dias et al. 2011, Fifield et al. 2014, McFarlane Tranquilla et al. 2014, Yamamoto et al. 2014, Lea et al. 2015). Distances can also be calculated between pairs of migratory routes (e.g. for the same individual during consecutive years), either between positions at certain landmarks (Yamamoto et al. 2014), or as the mean distance between each position on one route and the nearest position on the other (Guilford et al. 2011, Dias et al. 2013). The smaller that distance, the more consistent the individual. To compare the within- and between-individual distances (i.e. evaluate the statistical significance of individual consistency), one approach is to use distances calculated for pairs of centroids or routes as the response variable in a (G)LMM with individual (same vs. different) as a random effect, and check for the significance of the random effect (Dias et al. 2013). The second, and more widely used approach, is to compare the calculated within-individual distances with a null distribution of distances generated by reshuffling either locations or migratory tracks between individuals (Navarro & González-Solis 2009, Dias et al. 2011, Fifield et al. 2014, McFarlane Tranquilla et al. 2014). This method based on distances does not take into account the spread of the locations around the centroids: hence, although useful to detect a shift in the general distribution, it would not detect a change only in range size. It also has the disadvantage of only allowing individual specialisation to be detected, but not quantified.

Spatial analyses: overlap between distributions

Specialisation can also be estimated as the overlap between distributions of the same individual over
time. One approach is to overlay the locations (dives, landings or feeding events etc.) on a grid, and count the number of shared grid cells between different trips made by the same individual (Hedd et al. 2001, Baylis et al. 2015b, Orben et al. 2015, Sommerfeld et al. 2015). These values are usually compared with null models based on randomisation of individual identities. Problems include the sensitivities to grid cell size and to the resolution of the tracking data. Indeed, if the data are too coarse, there is a risk that genuine differences between individuals will be missed. Ideally, data should be analysed on a scale that is as fine as possible, although not smaller than the accuracy of the tracking device; however, if the grid cells are too small, potentially no 2 points from the same bird will ever fall in the same cell even if these points are relatively close.

Probabilistic measures offer an alternative approach; a utilisation distribution (UD) is generated from tracking data, and the polygons representing core and general use areas (typically 50% and 90/95%, respectively) are then compared in an analysis of overlap to determine the probability of individuals being located repeatedly in the same area. This method has been used to compare foraging areas in consecutive trips during the breeding season (Phillips et al. 2006, Pettex et al. 2012, Soanes et al. 2013), and areas used from one year to the next (Chilvers 2008, McFarlane Tranquilla et al. 2014, Muller et al. 2014). The problem is that it does not exploit the information on the complete UDs (cf. Ceia et al. 2014, 2015, Fifield et al. 2014, Wakefield et al. 2015). In contrast, the indices described by Fieberg & Kocchany (2005) provide a more elegant means of representing the overlap between pairs of UDs, mainly based on the product of 2 UDs. They suggested the use of Bhattacharyya’s affinity when the aim is to quantify the degree of similarity among UD estimates (see Wakefield et al. 2015), and the utilisation distribution overlap index (UDOI) when a measure of space-use sharing is desired. Isopleths can, of course, still be informative when using these indices. The observed distribution of the indices can be compared to randomised distributions. In general for methods using UDs, care needs to be taken in the definition of the smoothing factor \( h \) value required for kernel analysis, since it can influence the resulting UD. A constant value of \( h \) for all individuals should be preferred (Fifield et al. 2014); otherwise, variation in behaviour can be indistinguishable from that due to the choice of smoothing parameter.

### Environmental (habitat) analyses

Tracking data provide information not only on the geographical space, but also on how individuals use their environment (habitat). Indeed, remotely-sensed environmental data can be extracted for each animal location and, as with other types of movement information, traditionally each environmental dimension is analysed separately, typically calculating the overlap (Bhattacharyya’s affinity) between the usage distributions represented by pairs of trips for each individual (Wakefield et al. 2015). Alternatively, it is possible to include all variables in the same model, using random slope models (allowing for the response to environmental conditions to vary between individuals). This has been used to investigate how environmental conditions influence the track characteristics, e.g. speed or straightness of the path of different individuals (Patrick et al. 2014). The same approach using random slopes can be used in a resource selection function framework in which habitat selection is estimated by contrasting environmental conditions at ‘used’ locations (i.e. the recorded locations) and ‘available’ locations (randomly simulated locations in the accessible area around the recorded locations), in general using a generalised linear model (GLM) with a logistic link. This can be extended to GLMMs, using random slopes (individuals as a random effect) to detect differences between individuals in selection for each variable.

### CONTROLLING FOR PSEUDO-REPEATABILITY

Some effects can create bias in the estimation or interpretation of levels of individual variation, which can inflate repeatability estimates, leading to pseudo-repeatability (Dingemanse & Dochtermann 2013). This inflation occurs when predictor variables (i.e. fixed effects) that influence within-individual variation vary between individuals because of a sampling or measurement error, or when biologically relevant parameters (i.e. fixed effects) that explain between-individual differences are not taken into account in models (Dingemanse & Dochtermann 2013, Westneat et al. 2015). Other explanations proposed for heterogeneity in residual within-individual variance are the ‘organismal error’ (when the variance in phenotype is due to errors made by individuals when assessing their environment, e.g. individuals misidentify an environmental cue and produce a response which would be better suited to another environment) and the ‘random residual within-individual
Several issues related to temporal scale exist. Indeed, consistency detected at different timescales has different ecological interpretations (see Réale & Dingemanse 2001 for a related discussion on the study of animal personality). If individuals are consistent over a short but not over a long timescale, the ‘specialisation’ detected is likely to be due to variation in the state of the individuals (e.g. hunger level or reproductive state), or other short-term uncontrolled effects (e.g. immediate environmental conditions). If individuals are consistent over a long timescale, the cause is likely to be due to genetic, parental, individual quality or possibly permanent environmental effects. If specialisation increases over long timescales, the causes are likely to be related to some learning process (over the lifespan of an individual) or selective disappearance (over several generations, i.e. if specialists are fitter, generalists will selectively disappear from the population). This emphasises the importance of carrying out studies that, ideally, cover multiple time intervals (Kernaleguen et al. 2015).

Conversely, incorrect combination of time periods can lead to erroneous interpretations. Indeed, if individuals specialise on different resources or environments over different seasons, studying specialisation over the whole year, for example, might prevent the detection of individual specialisation (although this issue can be dealt with by including the correct fixed effects). In the case of seabirds, although several studies have revealed that dietary and behavioural specialisations are widespread, it is unclear for how long these specialisations are maintained (Masello et al. 2013, Patrick et al. 2014; but see Wakefield et al. 2015). It is likely, however, that repeatability in foraging behaviour declines at longer temporal scales because of temporal changes in the availability and predictability of resources (Woo et al. 2008, Bell et al. 2009, Ceia et al. 2014). Weimerskirch (2007) also suggested that site fidelity not only depends on the timescale, but also on the habitat visited. Almost all published studies to date were limited to data from relatively few individuals tracked or observed over short periods of time (Żydelis et al. 2011). Novak & Tinker (2015) also raise this point for time-aggregated observations related to diet, noting that increasing sampling time increases knowledge of an individual’s diet but comes with the risk that the ability to detect meaningful temporal patterns in prey selection is reduced.

The timing of the study can also bias the results for reasons that are not necessarily linked to individual preferences. For example, if individuals are tracked only during consecutive trips or for a few consecutive years, when conditions may be more similar than after longer intervals, the lack of within-individual flexibility in behaviour may reflect either that there was no environmental change influencing prey availability, or site fidelity (Chilvers 2008, Pettex et al. 2012, Carneiro et al. 2016). However, if tracked for multiple years with contrasting environmental conditions and prey availability, the repeated use of an area would indicate site fidelity. Such consistency in space use could also reflect behavioural plasticity, and the two mechanisms are hard to tease apart.

Timing and spatial accessibility are also important when studying between-individual differences. If all individuals in the study do not have access to the same environment or prey (either because the conditions change or because individuals live in distant areas with different characteristics), then the population might exhibit apparent specialisation, even though individuals are not specialised. Thus, to be sure that it is indeed specialisation that is observed, studies should minimise the risk of differences in habitat or resource availability between individuals, which in practice is a major challenge. Note that this can to some extent be resolved by combining trophic markers such as stable isotope ratios with tracking in multiple years (Baylis et al. 2015b).

Finally, populations can exhibit different foraging behaviours depending on the season (e.g. for seabirds, even within the breeding season, energy requirements and the constraint of the colony usually differ between pre-laying, incubation, brood-guard and later chick-rearing). It is important to take these changes into account, otherwise apparent individual differences might arise as an artefact of mismatches in the temporal scale of the measurements.

**EFFECTS SPECIFIC TO STABLE ISOTOPE DATA**

Studies using stable isotopes as a measure of individual specialisation/consistency in resource use should ideally use methods that allow the partitioning of isotope variation between different factors, and individual effects, such as calculating the adjusted repeatability from mixed-effects models. It is also possible to control for some of this variation during sample collection and preparation. Factors that should be taken into account when estimating short- and long-term spatial consistency or dietary speciali-
sation ($\delta^{13}C$ and $\delta^{15}N$ values) relate to (1) the comparison of tissue types with different enrichment factors for carbon and nitrogen isotopes (Cherel & Hobson 2007, Quillfeldt et al. 2008), (2) the use of tissues with different turnover rates (e.g. plasma and red blood cells), but with some overlap in terms of diet integration period (Votier et al. 2010, Ceia et al. 2015), (3) the differences in baseline isotope ratios in foraging areas used by individuals (in the marine environment, different oceanic processes and sources of organic matter can result in spatial changes in baseline stable isotope ratios; Moreno et al. 2011), and (4) the variation in baseline isotope ratios between different periods and years (Araújo et al. 2011, Wakefield et al. 2015). In addition, intrinsic factors, which are linked to physiological and life history traits (sex, breeding stage, experience, reproductive status etc.) can also be taken into account, although this will tend to decrease repeatability estimates (see Wilson 2008 for a discussion).

**Tissue type**

Keratinous tissues such as feathers, fur, vibrissae and chelonian shells are enriched in $\delta^{13}C$ and $\delta^{15}N$ when compared with blood, even when synthesised over the same time periods, due to different protein sources, use of endogenous reserves during feather synthesis, or when plasma contains $\delta^{15}N$-depleted uric acid (Hobson et al. 1996, Cherel et al. 2005, Cherel & Hobson 2007, Quillfeldt et al. 2008). Lipid concentrations can also lead to particularly depleted $\delta^{13}C$ values (Bearhop et al. 2000, Votier et al. 2010). Lipid extraction from fatty tissues is therefore recommended prior to $\delta^{13}C$ analysis (Cherel & Hobson 2007, Wakefield et al. 2015). However, extraction techniques can affect $\delta^{15}N$ in an unpredictable manner (Cherel et al. 2005, Bond & Jones 2009, Wakefield et al. 2015). Ideally, 2 samples, one to measure $\delta^{13}C$ (delipidated) and one to measure $\delta^{15}N$ (non-delipidated), should be analysed (Paiva et al. 2010, Wakefield et al. 2015). The low lipid level of keratinous tissues, blood cells and of whole blood does not affect their $\delta^{13}C$, and so lipid extraction is not required (Cherel et al. 2005, Bond & Jones 2009, Matich et al. 2011, Ceia et al. 2012, 2015). High and varying concentrations of lipid in blood plasma, however, can result in depleted $\delta^{13}C$ values (Votier et al. 2010, Ceia et al. 2012, 2015). Lipid extraction can often be impracticable because of the small quantities of blood plasma (Votier et al. 2010, García-Tarrasón et al. 2015, Wakefield et al. 2015). In order to account for the remaining potential differences in enrichment factors, correction factors can be used (Cherel et al. 2005, Quillfeldt et al. 2008, Votier et al. 2010, García-Tarrasón et al. 2015) or tissue type included as a fixed effect in adjusted repeatability analysis (Wakefield et al. 2015).

**Overlap in diet integration periods**

Several studies have modelled short-term consistency in isotope ratios by comparing values between plasma and red blood cells collected in a single event (e.g. Ceia et al. 2012, 2014, 2015, Wakefield et al. 2015). Although each tissue has a different turnover rate, the integration of prey isotopes into body tissues is a continuous process, and the analysis of short-term consistency using the same blood sample inevitably leads to some overlap in the periods which the samples represent (Votier et al. 2010, Ceia et al. 2015). To overcome this issue, when combined with tracking analysis, some studies have collected blood in 2 sampling events associated with the capture (deployment) and recapture (retrieval) of tracking devices, using the red blood cell fraction from initial capture and the plasma fraction from the recapture for subsequent analysis (Votier et al. 2010, Ceia et al. 2015), or only the plasma collected during both events (García-Tarrasón et al. 2015).

**Spatial and temporal variation in isotopic baselines**

Marine isoscapes can change depending on nutrient source, primary productivity, depth, latitude and oceanic frontal region, which can confound direct comparisons of trophic levels between animals from different regions (Stowasser et al. 2012, Moreno et al. 2016). Information on isotopic ratios of potential prey from different foraging areas (which can be determined from tracking devices) are essential to distinguish the relative importance of prey vs. habitat specialisation (Moreno et al. 2016). Otherwise, it is not possible to tell whether a change in isotopic value from $t_1$ to $t_2$ (or tissue type $x$ and $y$) represents a wider diet or spatial niche, or the same diet consumed in areas or periods with different isotope baselines (Ceia et al. 2014, Moreno et al. 2016). The latter applies in particular to species that forage across environmental boundaries and change their foraging areas on a seasonal basis (Stowasser et al. 2012). However, because most marine organisms are associated with specific water masses, even when appar-
ent differences in trophic position are most likely related to the use of areas with different baselines, dietary differences may still be informative about the use of particular prey (Jaeger et al. 2014). To date, there has been only one study including estimates of baselines as fixed effects in adjusted repeatability analysis of δ^13C and δ^15N in the blood of northern gannets Morus bassanus from the Bass Rock, within and across years (Wakefield et al. 2015).

**EFFECTS SPECIFIC TO TRACKING DATA**

Studies of behavioural consistency and individual specialisation based on tracking data should, where possible, apply a variety of complementary approaches. It is straightforward to include fixed factors in analyses of summary statistics, but more difficult to control for these effects in spatial analysis. For the latter, the analyses have to be carried out separately for each level of those factors (e.g. males and females treated separately), but this implies the need for bigger sample sizes. In addition to sex and age effects, life-history stage and breeding status (failed or successful) can potentially influence foraging strategies and therefore should also be considered when testing for individual specialisation. Breeding status, for example, may not only affect timing of events (e.g. Croxall et al. 2005, Phillips et al. 2005, Bogdanova et al. 2011, Dias et al. 2011, Yamamoto et al. 2014), but also the use of migratory destinations and routes. When this information is not known, such effects can be reduced by restricting comparisons to particular time windows, for example to minimise the risk that apparent differences between individuals might relate to differences in migration schedules that ultimately took similar routes (Guilford et al. 2011). Behaviour may also differ between different types or phases of foraging trips (outward and return journeys, and hunting and searching for food), or migration (outward and return migration, and residence at the main staging and wintering areas). For birds, during chick-rearing, different types of trips can be performed by adults while provisioning their chicks; short trips to maximise delivery rate per unit of time and long trips that enable adults to restore their own reserves (Weimerskirch et al. 1994, Weimerskirch 1998; but see Phillips et al. 2009). All these potential differences need to be taken into consideration when analysing individual specialisation, as the constraint to return sooner to the colony can prevent individuals from visiting areas that would otherwise be optimal.

**CONCLUSIONS**

Individual specialisation can be calculated by using repeated measurements for each individual, then calculating the within- and between-individual variation. Although there are several alternatives (see Table 1), the most common and flexible approach is to calculate repeatability, using the variances extracted from GLMMs with individual as a random effect (either random intercepts or random slopes). To obtain estimates of individual specialisation, care needs to be taken to exclude effects that can lead to pseudo-repeatability. In addition, analysing a variety of data types simultaneously can provide better insights. Analysis and interpretation can be improved if the study involves a representative number of individuals having access to the same habitat and resources (preferably over the same periods), and over short and long timescales. Statistical analyses of individual differences should be rigorous and follow advice mentioned in this review. Studies that conformed to these recommendations have found convincing evidence of behavioural consistency and individual specialisation in marine predators, albeit typically over relatively short timescales, as well as in a wide range of other taxa (e.g. Woo et al. 2008, Matich et al. 2011, Ceia et al. 2014, Patrick et al. 2014, Wakefield et al. 2015). In contrast, the ecological implications of consistent differences in resource or habitat selection at the individual or population level remain unclear (Ceia & Ramos 2015). However, such differences are likely to affect the conclusions of population dynamics models (as does individual consumer behaviour or trait variation; Okuyama 2008, Schreiber et al. 2011). Indeed, individual specialisation can affect interactions between individuals (e.g. by reducing intraspecific competition; Bolnick et al. 2011, Matich et al. 2011). Such differences may also reduce the predictive power of existing ecological models, for example species distributions models that are used increasingly to predict the response of a species to climate change or following an introduction (Pearman et al. 2008), wherein projections are made from average values for the population. Indeed, if individuals differ in their environmental tolerance (or preferences), species–environment relationships inferred from only a sample of individuals might not be representative of the ability of the species to cope with change.

Although in this review we have presented the most common methods used by the marine ecology research community to study individual specialisation and behavioural consistency, it is important to
note that there have been a number of interesting methodological developments in the animal personality field which build on repeatability analysis to ask targeted research questions, especially related to within- vs. between-individual variation, and partitioning of variance components (van de Pol & Wright 2009, Twiss & Franklin 2010, Dingemanse & Dochtermann 2013). Several of these methods have the potential to be applied to marine predators. Furthermore, the analysis of movement data has become increasingly sophisticated, with the development of state-space models and approaches based on hidden Markov models that can be applied to the detailed information from whole tracks. These relatively new approaches are statistically complex and have yet to be used in the context of individual specialisation, but offer the possibility of capturing fine-grained behavioural responses and preferences that are likely overlooked when summarising movements with a few simple statistics. Combined with the decreasing costs and increasing accuracy of biologging devices, we expect these new methods to greatly increase our ability to study specialisation in marine predators.

Acknowledgements. This study represents a contribution to the Ecosystems component of the British Antarctic Survey’s Polar Science for Planet Earth Programme, funded by The Natural Environment Research Council. We thank the anonymous referees for their many helpful suggestions for improvements to the paper.

LITERATURE CITED


Carneiro et al.: Individual specialisation in marine predators

163

Spectrum 29:2328–2336


Hobson KA, Schell DM, Renouf D, Noseworthy E (1996) Sta-


Staniland IJ, Robinson SL (2008) Segregation between the sexes: Antarctic fur seals, Arctocephalus gazella, foraging at South Georgia. Anim Behav 75:1581−1590


Thieman GW, Iverson SJ, Stirling I, Obbard ME (2011) Individual patterns of prey selection and dietary special-


Editorial responsibility: Jacob González-Solís (Guest Editor), Barcelona, Spain

Submitted: July 25, 2016; Accepted: May 31, 2017

Proofs received from author(s): July 8, 2017