Designing mate choice experiments

Liam R. Dougherty*

Institute of Integrative Biology, University of Liverpool, Liverpool, L69 7RB, UK

Page header: Designing mate choice experiments

*E-mail: liam.dougherty@liverpool.ac.uk; Tel.: (+44) 151 795 7771

ABSTRACT

The important role that mate choice plays in the lives of animals is matched by the large and active research field dedicated to studying it. Researchers work on a wide range of species and behaviours, and so the experimental approaches used to measure animal mate choice are highly variable. Importantly, these differences are often not purely cosmetic; they can strongly influence the measurement of choice, for example by varying the behaviour of animals during tests, the aspects of choice actually measured, and statistical power. Consideration of these effects are important when comparing results among studies using different types of test, or when using laboratory results to predict animal behaviour in natural populations. However, these effects have been underappreciated by the mate choice literature to date. I focus on five key experimental considerations that may influence choice: (1) should mating be allowed to occur, or should a proxy behavioural measure of preference be used instead? (2) Should subjects be given a choice of options? (3) Should each subject be tested more than once, either with the same or different stimuli? (4) When given a choice, how many options should the subject choose between? (5) What form should the experimental stimuli take? I discuss the practical advantages and disadvantages of common experimental approaches, and how they may influence the measurement of mate choice in systematic ways. Different approaches often influence the ability of animals to perceive and compare stimuli presented during tests, or the perceived costs and benefits of being choosy. Given that variation in the design of mate choice experiments is likely unavoidable, I emphasise that there is no single ‘correct’ approach to measuring choice across species, although ecological relevance is crucial if the aim is to understand how choice acts in natural populations. I also highlight the need for quantitative estimates of the sizes of potentially important effects, without which we cannot make informed design decisions.

Key words: mate choice, mating preference, experimental design, context-dependent, preference function, mate sampling, choice test, repeatability, comparative evaluation, cognition.

CONTENTS

I. INTRODUCTION

(1) The problem

Choosing the right mate is hugely important for all sexual animals. This is because mates often vary in fitness-related traits, so that there may be strong reproductive benefits to choosing some partners over others, because of the resources they can provide or the genes they will donate to offspring (Andersson, 1994). Accordingly, some form of mate choice has been observed in a wide range of animal groups (Andersson, 1994; Rosenthal, 2017), in males, females, or both (Johnstone, Reynolds & Deutsch, 1996; Bonduriansky, 2001; Edward & Chapman, 2011; Rosenthal, 2017). Mate choice is also an incredibly important evolutionary process, because it can influence which individuals successfully donate genes to the next generation, leading to evolution *via* the process of sexual selection (Darwin, 1871; Andersson, 1994). The important role that mate choice plays in the lives of animals, and in the evolution of populations, is matched by the large and active research field dedicated to studying it (Rosenthal, 2017). Mate choice can be defined as any behaviour which leads to an individual (the ‘chooser’) being more likely to mate with some members of the opposite sex than others (Halliday, 1983). We can study mate choice in a variety of ways. For example, a purely observational approach is to look for differential mating success in natural populations, and perhaps correlate this with traits possessed by the chosen sex. However, there are other processes besides mate choice (such as intrasexual competition; Andersson, 1994) which can lead to differential mating success, and which may be hard to rule out in correlational studies. A more common approach is to perform a mate choice experiment, either in the field or under laboratory conditions, in which we can control the identity of the chooser (hereafter I also refer to an individual used in a mate choice experiment, and whose behaviour we record, as a ‘subject’) and the mate options they are presented with. Importantly, mate choice experiments enable us to observe choices (the outcome of mating interactions) and also potentially measure mating preferences; that is, which traits or trait values are most preferred by choosers. Such preferences are an internal property of an individual, but with the right experimental design can be inferred from a subject’s behaviour or choices (see Section II.1).

However, while reading the mate choice literature one comes to a striking realisation: no two mate choice experiments are the same. Rather, the experimental protocols used to examine animal mate choice are highly variable across studies. For example, in some studies subjects are able to interact fully with mates, whereas in others they are presented with mating calls played from different speakers. In some cases each individual may be tested multiple times with different stimulus combinations, whereas in others each individual is only tested once. This variability in experimental design likely arises for three main reasons. First, mate choice itself is a highly diverse process, in terms of the species that express choice, the traits that are targeted by choosers (e.g. physiological, morphological, or behavioural), and the senses used to assess those traits (e.g. visual, acoustic, or chemical). This diversity means that a given experimental protocol may not be appropriate in every species, for practical or logistical reasons; we can't test whether female guppies prefer to mate with red-bellied males in the same way that we test whether male spiders prefer to mate with well-fed females. Second, the questions being asked vary. For example, experimenters may focus on different aspects of choice; from the sensory, neuro and cognitive processes leading to decisions to the evolutionary causes and consequences of choosing some mates over others. Third, mate choice is a complex process which occurs in natural populations, and distilling this complexity into an experimental setup can be difficult, given that our experiments also need to be statistically robust and logistically feasible. Juggling these competing interests leads to many difficult design choices and compromises.

Nevertheless, the differences in experimental design seen across mate choice experiments are often not simply cosmetic, but may significantly influence our measurement of choice. This happens for three main reasons. First, differences in experimental design may lead to fundamental differences in what is and isn't being tested. Second, they may influence the statistical power to detect significant effects. Third, differences in experimental design can influence how subjects behave during a test. There is good evidence to suggest that for many species choosers are constantly assessing their environment, and adapt their behaviour accordingly, and if we are not careful they may ignore the features of the experiment we want them to respond to, and instead respond to the features we want them to ignore. For these reasons many aspects of experimental design can significantly influence which traits are most preferred, how strong those preferences are, and even whether we see the expression of choice at all.

An example will help to illustrate both of these points. One fundamental aspect of mate choice experiments that can vary is the ‘choice design’: the number of options subjects are presented with. Choice can be measured either using a ‘choice test’, in which subjects are presented with two or more options simultaneously, or a ‘no-choice test’, in which subjects are presented with a single option and have to choose between this option or nothing (Wagner, 1998; Section IV.1). Results from these different designs may be compared to each other, which is problematic in two ways. First, these two designs test fundamentally different types of mate choice (Wagner, 1998). Should we expect the outcomes of these different tests to be the same, or might choice look different depending on which design we use? Second, unavoidable differences in experimental setup between the two designs may cause animals to behave differently. For example, choice tests may facilitate choice by making it easier for subjects to compare directly options presented simultaneously (Wagner, 1998). Notably, a meta-analysis of studies that tested mating preferences using both designs found that preferences are significantly stronger in choice then no-choice tests (Dougherty & Shuker, 2015*b*; Section IV.2).

This example illustrates how a single fundamental aspect of experimental design may influence both what is being tested, and how animals behave, during a mate choice experiment. Importantly, this design consideration may lead to systematic, non-random differences in the measurement of mate choice. This is a problem when we compare results from multiple studies, as differences in behaviour may be partly driven by experimental, rather than biological, factors. As in the case of choice designs, studies are often compared without consideration of these potential differences, and until recently there has been a lack of empirical tests of their magnitude. There is now a growing realisation that ignoring these effects has the potential to hinder our understanding of mate choice.

(2) Review outline

In this review I show how multiple fundamental aspects of experimental design can influence the measurement of mate choice, either by influencing chooser behaviour or our ability to measure that behaviour, and are often ignored when comparing empirical results across studies. Since the influential review by Wagner (1998), several of these issues have been discussed elsewhere (e.g. Hutchinson, 2005; Powell & Rosenthal, 2016; Lahti, 2015; Rosenthal, 2017). However, given the rapid development of the field and the large literature, a focused and updated review should be valuable in expanding and clarifying these issues.

I begin by briefly introducing how we typically measure mate choice. I then discuss how the expression of mate choice in animals is highly context dependent, and may be influenced by the state of the chooser, or by a range of social, demographic, and environmental factors which may be inadvertently altered during experiments. I then discuss five key design considerations that need to be made when planning a mate choice experiment. The five considerations are: (1) should mating be allowed to occur, or should a proxy behavioural measure of preference be used instead? (2) Should subjects be given a choice of options? (3) Should each subject be tested more than once, either with the same or different stimuli? (4) When given a choice, how many options should the subject have to choose between? (5) What form should the experimental stimuli take (for example, should subjects be presented with live animals, natural signals or artificial signals)? For each design consideration I will discuss alternative experimental approaches. I examine the advantages and disadvantages of each approach, and consider how different approaches may influence either how animals behave or what aspects of choice we are measuring. Where possible I will also discuss practical and statistical considerations of different approaches. Table 1 provides an overview of the main advantages and disadvantages associated with each approach. Finally, I suggest future research directions that would allow us to quantify, and potentially overcome, these problems.

I focus here on experimental design issues that are specific to mate choice; for this reason I do not cover important design considerations that apply to studies of animal behaviour more generally, such as appropriate randomisation, the use of proper controls, and the problems associated with measuring different types of behaviour (see Martin & Bateson, 2007). However, one general issue that mate choice experiments may be especially vulnerable to is low statistical power, and I discuss ways to mitigate this where possible. My intention is to encourage researchers to think carefully about the design decisions they make when planning a mate choice experiment, and to provide a framework for informing these decisions. While I define five main design considerations here, each cannot be considered in isolation: decisions relating to any single aspect of design may influence the advantages and disadvantages of other approaches, and specific logistical or biological issues may be mitigated or exacerbated by multiple design choices. The considerations discussed in Sections IV, V and VI are linked in this respect. Additionally, although researchers tend to focus on laboratory studies of mate choice, many of these problems also apply to measurements of choice in the field.

Importantly, I emphasise that there is no single ‘correct’ approach to take when designing mate choice experiments, rather that different approaches are suitable in different situations. Further, mate choice experiments often have different goals that influence the extent to which these experimental issues are confounding factors that need to be controlled. For example, studies may aim to quantify preferences under ideal conditions or as they are likely to occur in nature. In the latter case, experimental design differences are an unavoidable consequence of testing choosers in ‘realistic’, species-specific conditions. My key premise is that, when comparing across studies, we need to recognise that differences in choice behaviour can potentially arise purely due to differences in experimental design, rather than species differences in biological or environmental variables.

II. MATE CHOICE IN CONTEXT

(1) Mate choice theory and terminology

Mate choice is traditionally defined as occurring whenever the effects of traits expressed in one sex lead to non-random mating with members of the opposite sex (Halliday, 1983; Kokko *et al*., 2003; Edward, 2015). More recent definitions broaden this to include mating with same-sex individuals, which may have evolutionarily important consequences in some species (Rosenthal, 2017). Importantly, mate choice is an outcome as well as a process (Wiley & Poston, 1996; Jennions & Petrie, 1997). In other words, choice can only be said to have occurred after a mating has taken place. This is important because individuals are assumed to have internal, idealised mating preferences that become realised as choices (Jennions & Petrie, 1997). Because we cannot see into the heads of choosers, we infer these preferences from the choices they make (Rosenthal, 2017). However, it is worth bearing in mind that the distinction between mate choice and mating preference is not always made clear in the mate choice literature, and the terms ‘choice’ and ‘preference’ are often used interchangeably (Rosenthal, 2017). In general I also continue to use the term preference here as a shorthand when referring to the results of mate choice experiments. Importantly, authors may refer to experiments as ‘mate choice experiments’ even if the experiment does not directly measure mating (Section III.2), and I continue this convention here for convenience.

Mating preferences can be described using a ‘preference function’, which describes how the likelihood of accepting a mate changes in relation to a mate phenotype (Lande, 1981; Kirkpatrick, 1982; Basolo, 1990*a*; Wagner, 1998; Kilmer *et al*., 2017). Preference functions may be linear and directional, or complex and non-linear (Ritchie, 1996; Edward, 2015). Irrespective of shape, the two components of a preference function that are most likely to vary are: the most preferred trait (the peak preference) and preference strength (Fig. 1). The peak preference is the trait value that elicits the greatest response in a chooser (Reinhold & Schielzeth, 2015; Edward, 2015). The strength of a preference (also referred to as choosiness or selectivity: Edward, 2015) describes how strongly a given trait value is preferred in relation to other values (Reinhold & Schielzeth, 2015). Here I consider a choosy or selective individual to be one who shows a strong difference in response to preferred over non-preferred trait values. This is reflected, for example, in a steeper slope when applied to linear or quadratic preference functions (Fig. 1). Note that this use of ‘choosiness’ is different from the other common definition used in the mate choice literature (the effort an individual is prepared to invest in mate assessment: Jennions & Petrie, 1997).

(2) Mate choice is highly context dependent

A key facet of mate choice is that it is strongly dependent on the environment, both in terms of which traits are preferred and the extent to which preferences can be realised. Importantly, the fact that choice is an outcome means that the expression of preferences may often be constrained. Preferences can be constrained most simply by the options available at any one time; individuals cannot choose the most preferred mate if all of the available options are poor. Further, if animals are under pressure to mate quickly, or typically have few opportunities to mate in the wild, then they may settle for less-attractive options to ensure successful reproduction (Jennions & Petrie, 1997; Ah-King & Gowaty, 2016). Choosers in such environments may thus appear to have different preferences to those in environments with a wide range of available mate options. Similarly, changes in sampling tactics have the potential to alter choice outcomes, and may influence all aspects of preference functions, including preference strength and peak preference (Jennions & Petrie, 1997; Rosenthal, 2017). The environment may also often constrain the ability of animals to sample mates, for example when the travel cost between mate options is high (e.g. Magnhagen, 1991; Milinski & Bakker, 1992; Booksmythe, Detto & Backwell, 2008). Mate choice is also constrained because mate sampling and assessment are cognitive processes (Ryan, Akre & Kirkpatrick, 2009). Any factors that reduce the ability of animals to perceive mates and make decisions will constrain realised choice (Section VII). For example, the ability to assess acoustic stimuli is reduced in noisy environments (Swaddle & Page, 2007; Bee & Schwartz, 2009), and visual cues are harder to assess when visibility is poor (e.g. Seehausen, Van Alphen & Witte, 1997; Candolin, Salesto & Evers, 2007). In these cases impairments in the ability to perceive or compare stimuli may reduce choice accuracy – the ability to identify the highest quality option. Finally, choice can also be constrained by the behaviour of mates. For example, males may force matings (Shuker & Day, 2001; Arnqvist & Rowe, 2005), and both sexes may disrupt courtship by rivals (Wong & Candolin, 2005; Baxter *et al*., 2018) and aggressively guard mates after mating (Simmons, 2001).

As well as varying non-adaptively due to environmental constraints, mate choice can also vary adaptively if different mate phenotypes are selected in different environments. For example, spadefoot toad (*Spea bombifrons*) females show a preference for heterospecific males only in environments in which hybrids do better than non-hybrids (Pfennig, 2007). The expression of choice may also vary in relation to the costs and benefits of being choosy. While mate choice may provide choosers with direct and indirect fitness benefits (Andersson, 1994; Kokko *et al*., 2003), it can also entail costs, such as the time, energy or predation costs associated with sampling mates (Magnhagen, 1991; Sullivan, 1994; Byers *et al*., 2005; Vitousek *et al*., 2007; Hughes, Kelley & Banks, 2012), or the risk of remaining unmated if preferred mates are not available (Werner & Lotem, 2006; Barry & Kokko, 2010; Greenway, Dougherty & Shuker, 2015). The net benefit of being choosy can vary strongly in different contexts. For example, in areas of higher mate density the cost of sampling is lower and the potential number of mating opportunities is greater. When mate density is higher choosers may therefore exhibit stronger mating preferences (e.g. Berglund, 1995; Kokko & Rankin, 2006; Svensson, Lehtonen & Wong, 2010). Animals frequently exhibit plasticity in the strength of their mating preferences in response to changes in demographic and environmental factors that influence the costs and benefits of mate choice, including the population density or mate encounter rate and the operational sex ratio (Jennions & Petrie, 1997; Ah-King & Gowaty, 2016; Kelly, 2018). Additionally, the costs of choice may be influenced by the state of the chooser (Cotton, Small & Pomiankowski, 2006). For example, choosers in poor condition may have an above average mortality rate, and therefore fewer potential opportunities to mate (Cotton *et al*., 2006). Poor-condition individuals may then become less choosy to ensure successful reproduction before death (e.g. Hingle, Fowler & Pomiankowski, 2001; Hunt, Brooks & Jennions, 2005), or choosier if mates provide direct benefits which can improve survival (Wagner & Harper, 2003).

Finally, mate choice is ultimately a social process that can be influenced by the social information available either before or during the choice process (Rodríguez, Rebar & Fowler-Finn, 2013*b*). Most commonly, the social environment influences which traits are most preferred by choosers. For example, a range of species engage in mate choice copying, whereby individuals use social information from same-sex rivals to inform their own choices (e.g. Dugatkin, 1992; Vakirtzis, 2011). In many species preferences are frequency dependent, with common or recently experienced phenotypes being preferred in some cases (e.g. Miller & Fincke, 1999; van Gossum, Stoks & Bruyn, 2001; Hebets, 2003), and rare or novel phenotypes being preferred in others (e.g. Zajitschek & Brooks, 2008; Royle, Lindström & Metcalfe, 2008; Janif, Brooks & Dixson, 2014). Additionally, mating preferences are sometimes altered by the types of mates or stimuli encountered during development (e.g. Hebets, 2003; Bailey & Zuk, 2008; Jordan & Brooks, 2012). Indeed, in many bird species mating preferences are partly learned during a sensitive period of development (Immelman, 1975; ten Cate & Vos, 1999; Slagsvold *et al*., 2002; Grant & Grant, 2018).

In summary, the choices that animals make are inextricably linked to their state, their prior experience, and the environments in which they are tested. A chooser’s state or environment has the potential to influence several components of choice, including: the traits used to assess mates, the strength of preferences, the traits values that are most preferred, and the sampling strategies used to assess mates. Small changes in experimental design therefore have the potential to alter the expression of mate choice, and often in subtle ways. For example, differences in the rearing environment between experimental groups could unintentionally affect the condition of subjects, or the sex ratio they experience, which may influence their subsequent mating behaviour. The social environment especially may be changed unintentionally, affecting the strength of choice. While the magnitudes of these effects are not clear for most species, in some cases they can be large, and often systematically alter choice behaviour. How might we account for these effects? One way is to design our experiments to exclude any confounding effects in which we are not interested, for example, by rearing subjects in social isolation to exclude experience effects. However, as we will see, many basic design decisions may unavoidably influence some of these effects.

Another is to measure choices or preferences in several contexts, and then quantify the extent to which they change. Both of these solutions are limited in practice: no experiment can remove all confounding effects, and we cannot perform the same experiment in every conceivable environment. Instead we have to be pragmatic, and temper our conclusions accordingly. For this reason, I encourage experimenters to be explicit about the context of their experiments, so that others can judge the generality of their findings. For example, one might report that: 'Our results suggest that young virgin females, reared in social isolation, and maintained on an *ad-libitum* diet prefer…'. This will allow others to understand more clearly the context of a result, and will facilitate comparison across studies.

III. CHOICES *VERSUS* PREFERENCES

One of the first decisions to be made when designing a mate choice experiment is: should one record actual mating events, or another behavioural measure of preference? Measuring choice outcomes requires males and females to interact physically and mate. Alternatively, one can record a behaviour that is assumed to correlate with a mating preference, known as a 'proxy' measure of preference. For example, approach latency towards a speaker is the most common measure of preference in amphibians in which mates are attracted with long-distance song (Wagner, 1998). These are termed proxy measures because mating preferences are not measured directly: we simply infer what an animal 'prefers' from its behaviour (Section II). Each of these approaches has its advantages and disadvantages, which I will discuss in turn.

(1) Mating outcomes

Mating is required for mate choice in the strict sense, because only then does choice influence reproductive fitness. Therefore researchers should allow behavioural interactions to culminate in mating if they are interested in the potential evolutionary consequences of mate choice. This includes studies of the strength and shape of sexual selection acting on traits in the choosing or chosen sex, or how mate choice influences hybridisation between species (e.g. Coyne, Elwyn & Rolán-Alvarez, 2005; Chenoweth & Blows, 2005; Dougherty & Shuker, 2014). Additionally, in some species mating may be the only reliable measure of choice. For example, many species use olfactory, chemical or tactile cues during courtship which require direct contact between mates before choosers can make an assessment (e.g. Spieth, 1974; Tallamy, Powell & McClafferty, 2002; Chenoweth & Blows, 2005), or may frequently reject potential mates in the latter stages of courtship or after genital contact has already been achieved (e.g. Simmons & Bailey, 1990; Pizzari & Birkhead, 2000; Tallamy *et al*., 2002).

However, allowing mating creates limitations. First, mating may influence future receptivity, so that it may not be possible to test subjects repeatedly (Rosenthal, 2017; Section V.2). Second, it is generally harder to manipulate traits in live animals, so it may be more difficult to establish a causal link between specific traits and chooser responses when allowing individuals to mate (Martin & Bateson, 2007; Section VII.1). Third, successful mating in internally fertilising species requires that males and females interact physically. This leads to the complication that male and female preferences may not co-align (Gavrilets, Arnqvist & Friberg, 2001; Kokko *et al*., 2003; Arnqvist & Rowe, 2005), and it may be difficult to ascertain who has ‘chosen’ whom. While in many species one sex takes on the role of courter and one as chooser, some form of mutual mate choice appears to be common, regardless of traditional sex roles (Sandvik, Rosenqvist & Berglund, 2000; Rosenthal, 2017). Further, because such behavioural interactions may be subtle and difficult to observe, experiments in which individuals are able to interact are unsuitable if we wish to attribute choice explicitly to either party (Halliday, 1983; Martel & Boivin, 2011). Additionally, it may be difficult to separate active choice from other, more indirect mechanisms of choice using such tests. For example, forced matings by males are common in many species, so that choice outcomes may depend on male persistence and female resistance behaviours (Gavrilets *et al*., 2001; Arnqvist & Rowe, 2005). In such species it is then not possible to measure strict female mating preferences when choice is inferred from mating (Shuker & Day, 2001; Baxter *et al*., 2018). Both of these problems also apply to measures of preference that use quantitative aspects of mating behaviour, such as copulation duration or amount of sperm transfer (Rosenthal, 2017). These measures have the added complication that post-copulatory investment (such as the number of sperm transferred during mating) may be driven by factors unrelated to mate choice, such as intrasexual competition (Simmons, 2001).

A related problem arises during choice tests in which subjects are presented with multiple mates simultaneously (Section IV). If rivals interact directly during tests, it may be hard to separate active choice from the effects of intrasexual competition (Andersson, 1994; Wong & Candolin, 2005; Baxter *et al*., 2018). One way this problem can be overcome is by giving subjects free access to both mates, but preventing mates from interacting directly with each other. This can be done by tethering mates while allowing subjects to have free movement (e.g. Pilakouta & Smiseth, 2017), or separating mates into compartments accessed by holes that only the choosing sex can pass through (e.g. Baxter *et al*., 2018).

(2) Proxy measures of preference

The mate choice literature uses a wide range of proxy measures of mate preference (Rosenthal, 2017), including the time spent associating with or interacting with stimuli, the latency to approach stimuli, solicitation behaviours, positive responses to courtship, the number of choices in an operant task (e.g. Holveck & Riebel, 2007), the duration of attention (determined using eye-tracking technology: e.g. Dixson *et al*., 2011; Yorzinski *et al*., 2013) or courtship attempts directed at mates. Non-behavioural approaches are also used to determine mating preferences, particularly in humans. This includes measures of neural activity (e.g. Aharon *et al*., 2001; Cummings, 2015) and genital arousal (Chivers *et al*., 2010). Most proxy measures of preference are most applicable to species that use visual, acoustic or olfactory courtship traits that can be assessed without direct contact with mates; for example, association time is often tested with mates separated by transparent screens. As with any behavioural measurement, proxy measures of preference need to be defined robustly, otherwise subjective experimenter biases could influence behavioural measurements (Rosenthal, 2017). This problem can be minimised by the use of blind measurements (Martin & Bateson, 2007; Holman *et al*., 2015), although this is not always possible for practical reasons.

The use of proxy measure of preference brings several benefits. First, recording preferences in the absence of direct interaction (as in the latter three options) means that attributing choice to either sex is not a problem, and makes it easier to establish a causal link between specific traits and chooser responses. This means that in species where courters are choosy or are able to force matings, proxy behavioural measures may reflect preferences better than choice outcomes (Section III.1). Second, focusing on proxy behaviours allows for a broader range of experimental setups to be used, because full interaction between mates is not required. For example, proxy behaviours can be recorded after subjects are presented with: (1) mates they are able to interact with fully, (2) mates they are unable to contact, (3) natural stimuli in the absence of mates (e.g. chemical or acoustic cues), or (4) artificial/virtual stimuli in the absence of mates (including models or videos; see Section VII.1). Importantly, the use of proxy measures allows for the creation of novel stimuli, including those outside the natural range of trait values, or the manipulation of existing traits that may not be possible when using live mates (Section VII.1). Third, proxy measures only require individuals of one sex to be reared or kept in experimental facilities, thus reducing the overall number of animals needed for an experiment of a given sample size. This may reduce the time and resources needed to perform an experiment, and reduces the number of animals used during a test without sacrificing statistical power. Fourth, there may be statistical advantages to using a continuous, proxy measure of preference (such as mating latency) compared to a dichotomous choice measurement (e.g. mated or not) which typically requires a larger sample size to detect a given effect (Wagner, 1998; Reinhold & Schielzeth, 2015; Section VIII). Therefore, proxy measures of preference can be useful when the number of individuals available for testing is limited.

One disadvantage of this approach is that it may be difficult to determine whether a proxy behaviour actually reflects a mating preference. This is because there may be other, non-sexual reasons why animals interact with or are attracted to members of the opposite sex. For example, in the sailfin molly (*Poecilia latipinna*) individuals of both sexes prefer to associate with large fish, irrespective of sex, and females given a choice between a large female and a small male preferred to associate with the large female (Gabor, 1999). This suggests that association may sometimes reflect a purely social preference, for example because size-assortative shoaling lowers individual predation risk (Hoare *et al*., 2000). More subtly, many forms of sexual interest could vary due to differences in personality (David & Cezilly, 2011) or mate sampling strategies (Wagner, 1998) rather than preferences, so that longer periods of interaction do not necessarily correlate with the perceived attractiveness of the mate. In this case, relative time spent with some stimuli over others may be a better measure of preference, as this controls for the total time each subject spends interacting. Other species may find the sides of an arena (whether these are association zones or not) more attractive than empty space in the middle for safety reasons. These problems can potentially be minimised by using a behavioural measure of preference that is only ever expressed in a mating context (Kilmer *et al*., 2017; Rosenthal, 2017); these are known as ‘proceptive behaviours’ (Beach, 1976). Female phonotaxis in amphibians and copulation solicitation displays in birds are two suggested examples (e.g. Lynch *et al*., 2005; Amy *et al*., 2015). This may have the added advantage of reducing the likelihood that non-receptive individuals are included in analyses of choice (Rosenthal, 2017).

Importantly, even when proxy behavioural measures reflect mating preferences, they do not always translate into choices, for the reasons outlined in Section II.2. This means that, if one is interested in actual fitness outcomes, using a proxy behavioural measure is only valid if preferences have been confirmed to correlate with actual choice outcomes. Several studies in birds and fish have validated association preferences by matching them to mating frequencies or preferences observed in tests with freely interacting animals (Table 2). In other cases proxy behaviours do not match choice outcomes, perhaps because courters are also choosy or can force matings (e.g. Goncalves & Oliveira, 2003; Owen, Rohrer & Howard, 2012).

Another disadvantage of using a proxy measure of preference is that the flexibility mentioned above means there are potentially more design decisions to be made, relating to which behaviours are chosen and how they are measured. For animals with complex behavioural repertoires it may not always be obvious which behaviours best reflect preferences. Further, it is worth noting that different behaviours may reflect mating preferences to different extents. For example, studies that have examined preferences using several behavioural measures find mixed results: in some cases preferences are consistent across behaviours (e.g. Witte, 2006;

Cummings & Mollaghan, 2006; Holveck & Riebel, 2007; Jeswiet & Godin, 2011), whereas in others they are not (e.g. St John & Fuller, 2018). An additional consideration is how exactly preferences should be measured. For example, when recording association time, should subjects be looking at, or actively signalling to mates whilst in the association zone before timing begins? Even a minor factor such as the width of the association zone has the potential to influence the strength of preferences detected (Vukomanovic & Rodd, 2007). This highlights the importance of thinking carefully about all aspects of the experimental setup.

IV. CHOICE DESIGNS

An important way in which mate choice experiments can vary is whether or not subjects can choose between options during a test (the choice design or choice paradigm; Wagner, 1998; Dougherty & Shuker, 2015*b*). Experiments can use either a choice test or a no-choice test. In a choice test, each subject is presented with more than one option simultaneously. Most commonly, choice tests use two options (but see Section VI). In a no-choice test, each subject is presented with a single stimulus. Both no-choice and choice tests are commonly used: out of 127 papers citing the review by Wagner (1998) between 1998 and 2012, 72% used a choice test, while 23% used a no-choice test, and 5% used both (Owen *et al*., 2012). Both choice designs can be used to record either mating outcomes or proxy measures of preference. For example, latency to approach a speaker or mate is a common proxy measure of preference recorded during no-choice tests (Wagner, 1998). Additionally, mate choice may be measured following ‘group choice’ trials, in which multiple members of each sex are introduced into an arena, and mating outcomes are recorded (e.g. Coyne *et al*., 2005; Head, Lindholm & Brooks, 2008; Holveck, Gauthier & Nieberding, 2015). I will not consider these further here, although I note the problems associated with attributing choice, whilst ruling out intrasexual competition, in these tests (Section III.1).

(1) Design considerations

Subjects in each choice design are forced to make different types of choices. In a choice test, the subject is required to choose between (usually two) alternative options, whereas in a no-choice test the subject has to choose whether to respond at all to the available option. For this reason the term ‘no-choice’ is a misnomer: there is still a choice between accepting an option, or rejecting it and waiting for a future mating opportunity (Barry & Kokko, 2010; Rosenthal, 2017). Importantly, the absence of a positive response does not mean that a decision has not been made by the subject, although it may be hard to rule out motivation effects (I discuss non-responsive subjects later in this section). Because of this fundamental difference, no-choice tests measure absolute preferences for a target trait, whereas choice tests measure relative, directional preferences between traits (Wagner, 1998; MacLaren & Rowland, 2006). The two choice designs can be considered to test two fundamentally different aspects of cognition. Choice tests allow for the simultaneous assessment and comparison of options, whereas no-choice tests instead require that subjects compare a stimulus against either an internal template or a stimulus from memory. For this reason, no-choice tests are sometimes considered to test ‘recognition’ of the range of acceptable stimuli, whereas choice tests are assumed to test ‘discrimination’ (Phelps, Rand & Ryan, 2006; Ryan & Taylor, 2015). This also means that common cognitive or perceptual biases associated with comparing stimuli generally only apply to choice tests (Akre *et al*., 2011; Section VII.3).

Both types of choice design bring advantages and disadvantages. No-choice tests are especially useful when recording mating outcomes, as intrasexual competition is not possible.

However, several consecutive no-choice tests (often referred to as ‘sequential choice’ tests) are needed to examine a subject’s preferences across multiple stimuli, and multiple testing of the same individual can introduce additional problems (Section V.2). Indeed, no-choice tests are most suitable for testing preferences for signals that are assessed sequentially, such as contact pheromones; species that use sequential comparison may not even ‘see’ that other choices are available when presented simultaneously (Dougherty & Shuker, 2014; I will return to the issue of ecological realism in the next section).

For simple tests involving few options, choice tests reduce the number of trials that need to be performed on each subject compared to sequential tests, simply because choice trials test at least twice as many options as no-choice trials (Wagner, 1998; Section V). However, choice tests assume that subjects are able to perceive and assess multiple stimuli simultaneously. This assumption may be less realistic as the number of options increases due to cognitive or perceptual limitations (Section VI). Experimenters also need to rule out potential side preferences seen during dichotomous choice tests, by ensuring the sides symmetric or alternating sides between tests (Rosenthal, 2017). One potential problem with some choice tests is the assumption that subjects are always actively choosing between options; this can be mitigated by the use of neutral zones (which are standard for studies of association time: Section III.2) so that subjects have the option not to choose during the test (Lafleur, Lozano & Sclafani, 1997; Rosenthal, 2017). Further, because choice tests measure relative preferences, it has been suggested that two-choice tests could potentially exaggerate the strength of choice when used in combination with a dichotomous scoring system such as mating success, as even a small difference in preference is then constrained to an all-or-nothing response (Wagner, 1998; Fig. 2). This problem can be mitigated by using a continuous measure of preference (Section III.2).

Another difference between the two choice designs is how non-responsive subjects are treated. Subjects that do not choose in choice tests are usually excluded from further analysis, as they are assumed to be sexually non-receptive (e.g. Allison & Cardé, 2008; Beckers & Wagner, 2011). This is in contrast to no-choice tests, in which not choosing (failing to respond) is classed as a rejection and is integral to the analysis. It has been suggested that this biased inclusion of non-receptive individuals may lead to mating preferences being underestimated when using no-choice tests (Kokko & Jennions, 2015; but see Section IV.2). Crucially, there are alternative explanations for why subjects might not respond to stimuli, besides being non-receptive, which can make the exclusion of these individuals problematic. For example, total rejection of a given set of options could suggest that these options fall below a preference threshold (Valone *et al*., 1996). Further, rejection may be a common decision in species where mating is costly (Rosenthal, 2017), and giving subjects the option not to respond to any options during trials may more closely reflect choice in the wild, where animals vary in their motivation to mate (Phelps *et al*., 2006). One way to distinguish non-receptive from non-responsive subjects is to test subject with multiple options, or by using a control stimulus between tests that is known to elicit a strong response (Rosenthal, 2017; Section V.2). Unfortunately, the number of non-receptive individuals excluded during experiments is often not reported (Kokko & Jennions, 2015; Rosenthal, 2017). Given that this information is potentially informative, it should be reported in the methods section of studies as standard (e.g. Berglund, 1995; Bailey, 2011; Hayes *et al*., 2016).

(2) Differences in the strength of choice

Many studies have measured mating preferences using both no-choice and choice tests in the same species, for the same traits, and often find that preferences are stronger in choice tests

(Table 3). The magnitude of this effect was investigated in a phylogenetically controlled meta-analysis examining 38 of these studies (across 40 species), which showed that female (but not male) mating preferences are significantly stronger in choice than no-choice tests (Dougherty & Shuker, 2015*b*). Interestingly, other studies have shown that this effect is not limited to mate choice: similar differences between choice designs have also been seen for host choice in parasitoids and predators (Withers & Mansfield, 2005), foraging choices in hummingbirds (Morgan *et al.*, 2016), and virtual prey choice in humans (Beatty & Franks, 2012).

There are several methodological or behavioural differences between designs that could be driving this effect in relation to mate choice, although subsequent analysis of the 38 studies in the above analysis appears to rule out bias due to the exclusion of non-receptive subjects during choice tests (Dougherty & Shuker, 2015*a*). Instead, I suggest this effect is driven primarily by differences in subject behaviour in each choice test. Different choice designs may influence subject behaviour in two main ways. First, no-choice and choice designs differ in whether options can be directly compared or not during a test. This could in theory influence the ability of animals to assess and compare options (Bateson & Healy, 2005; Beatty & Franks, 2012). However, this depends on the mechanism of mate assessment used by subjects: species that rely on comparison of multiple options simultaneously during mate assessment may find it more difficult to assess options sequentially (e.g. Beatty & Franks, 2012), but species that rely on simple, threshold-based decision rules, or assess mates using contact-based cues (such as cuticular hydrocarbons in insects), may always assess options sequentially, irrespective of how many are available simultaneously (e.g. Kacelnik *et al.*, 2011; Dougherty & Shuker, 2014). Therefore this effect would only explain the above results if most of the species involved typically compare mates simultaneously, which is currently unclear.

Choice designs also differ in the perceived cost of rejecting an option, because of the risk that no other mating opportunities will arise. This risk arises because the likelihood of being presented with another option is unknown to subjects during an experiment (Valone *et al*., 1996; Section II.2). During a no-choice test the decision to reject the presented option has to be weighed against this perceived risk, whereas in choice tests the subject can reject one mate option, safe in the knowledge that there is another available. Therefore, subjects in a choice test only need to pick whichever option they prefer, even if the difference between the two options is very small. If subjects tested using a no-choice design perceive that the risk of remaining unmated is higher than in a choice design, then they will be more likely to mate randomly (De Jong & Sabelis, 1991; Werner & Lotem, 2006; Barry & Kokko, 2010), or respond maximally to all options (Rosenthal, 2017; Fig. 2). Importantly, subjects may use information obtained either through their current environment or rearing experience to predict the likelihood of remating. Therefore this problem may be exacerbated by the rearing conditions of animals prior to choice trials: subjects kept in same-sex rearing conditions, with no experience of mates, will have a very low perceived mate encounter rate, and may mate randomly during no-choice tests. A further consequence of this extra cost of rejection could be that only preferences above a certain threshold strength are expressed in no-choice experiments. In other words, we may be less likely to observe weak mating preferences in no-choice experiments compared to choice experiments, as there is no cost of choice in the latter situation. More generally, the sex ratio during mating interactions differs between no-choice tests (1:1 sex ratio) and choice tests (2:1 sex ratio or more, with the subject as the rarer sex), and population sex ratio influences the strength of mating preferences across a broad taxonomic range (Section II.2). In other words, two aspects of the social/demographic environment (the perceived mate encounter rate and population sex ratio) may be confounded between choice designs, with the potential to influence chooser behaviour.

The meta-analysis of Dougherty & Shuker (2015*b*) is a rare example of the quantification of the impact of experimental design on the measurement of mate choice, and reveals that the choice design strongly and consistently influences mate choice across species. The process, or combination of processes, driving this difference in behaviour remains unclear. Nevertheless, this quantification allows us to estimate the extent to which it is valid to compare choice estimates from studies that differ only in which choice design is used. Importantly, for studies interested in estimating choice as it occurs in natural populations, it also allows us to estimate the extent to which experiments may under- or overestimate the strength of choice, if they use a design that is inappropriate for their focal species. For example, the strength of choice may be significantly overestimated if choice tests are performed using species which typically encounter mates sequentially (Barry, Holwell & Herberstein, 2010; Booksmythe, Jennions & Backwell, 2011). This further supports the assertion that the choice design used should be informed whenever possible by the patterns of mate encounter most commonly seen in the wild (Dougherty & Shuker, 2015*b*). For most species we may have only a rough idea of the dynamics of mate encounter in natural populations, or how mates are actually perceived or assessed (see Sections VI.1 and VII.2). Field estimates of these parameters should be a priority if we wish to link mating preferences obtained in the laboratory to those in natural populations.

V. TESTING SUBJECTS MORE THAN ONCE

Another important methodological issue when designing mate choice experiments is how many times each subject should be tested. Importantly, testing subjects more than once has the potential to reveal within-individual variability in preferences.

(1) Reasons for multiple testing

From a logistical perspective, the simplest experimental approach is to test each subject once with either a single stimulus (no-choice test) or choice of stimuli (choice test). A preference for one trait value over another is then determined by taking the average response of all the subjects tested. This is known as a 'population-level' test (Wagner, 1998). One limitation of this approach is that it is unable to detect within-individual variation (or repeatability) in mating preferences. Such variation could arise because individuals show adaptive phenotypic plasticity in response to their state or the current environment (Section II.2), or because choice behaviour is inaccurate or inherently stochastic (Jennions, Kokko & Klug, 2012). Whatever the reason, within-individual variation in mate choice can be detected by testing each subject in several trials, to estimate repeatability (Widemo & Saether, 1999; Wagner, 1998; Brooks & Endler, 2001). Estimating the repeatability of choice allows us to infer the strength of recent selection on choice, because low preference variability can be indicative of strong recent selection that has depleted variation (Fowler-Finn & Rodríguez, 2013). An estimate of repeatability also allows us to estimate the potential future response to selection, because consistent preferences are needed to produce predictable evolutionary changes (Fowler-Finn & Rodríguez, 2013). Importantly, although within-individual repeatability in mating preference is often assumed to be high, surveys suggests that across species it is generally low (Bell, Hankison & Laskowski, 2009; Rosenthal, 2017), though the main factors driving this variability are not clear (Section II.2).

The second limitation of population-level tests is that they mask between-individual differences in preferences. Such variation may arise due to between-individual differences in preference alleles (Brooks & Endler, 2001; Rosenthal, 2017), state (Cotton *et al*., 2006; Section II.2), social experience (Rodríguez *et al*., 2013*b;* Section II.2), or sensory perception (e.g. Ronald, Fernández-Juricic & Lucas, 2012, 2018). This variation can be assessed by testing all subjects with the same stimuli, to determine 'individual-level' preferences (Wagner, 1998; Rodríguez *et al*., 2013*a*). The degree of variation in preferences can then be determined statistically, for example using mixed models which incorporate chooser ID as a random effect (e.g. Rodríguez *et al*., 2013*a*). Importantly, variation in preferences is harder to detect when each subject is tested with a limited number of options. Therefore, the more options each subject is tested with, the more variation can be captured (Roff & Fairbairn, 2015). For individual-level preferences, obtaining a high-resolution preference function may require multiple testing of each individual, depending on the choice design used during the experiment (Section IV). For no-choice tests, the only solution is to perform sequential tests on the same subject, with different stimuli presented each time. For choice tests, subjects can either be tested in multiple trials, with few options in each trial, or tested in a single trial with many options (although the latter approach means individual repeatability cannot be estimated). Importantly, both of these approaches have potential drawbacks (See sections V.2 and VI.2 respectively).

(2) Design considerations

While repeated testing may allow us to detect evolutionarily significant within- or between-subject variation in preferences, there are several additional problems that should be considered when using this approach. First, performing multiple trials per subject is logistically more difficult, and increases the time and cost needed to perform an experiment.

Importantly, this may reduce the statistical power to detect a preference if it leads to fewer subjects being tested with each stimulus (Raffa, Havill & Nordheim, 2002; Hutchinson, 2005). Time may also be an important limitation when testing species that have a short reproductive period.

Multiple testing may also alter the behaviour of subjects. In many species the strength of choice decreases as the interval between presentations increases (e.g. Bakker & Milinski, 1991; Svensson *et al*., 2010), probably because the perceived mate density decreases (Barry & Kokko, 2010; Booksmythe *et al*., 2011; Section II.2). Therefore, if possible, variation in the interval between presentations should be minimised to reduce this effect (unless one is interested in describing choice as it occurs in natural populations, for which mate encounter rate may be variable) . Additionally, receptivity or motivation to respond to a stimulus may decrease over successive trials (e.g. Pitcher *et al*., 2003), and this could potentially influence choice if not addressed. This may be especially problematic if tests allow for mating (Section III.1), which often influences subsequent receptivity to mates (Rosenthal, 2017). One way changes in receptivity can be assessed is by incorporating the time taken to make a choice (choice latency) into the statistical analysis (e.g. Bailey, 2008), except in those cases where this is itself used as a measure of preference. Alternatively, it may be possible to test receptivity using a control stimulus which is known to be attractive. For example, an experiment testing female responses to heterospecific calls could use a conspecific call as a control stimulus (e.g. Bush, Gerhardt & Schul, 2002). Response to this control stimulus can then be used to calibrate the response to the experimental stimuli.

Most importantly, repeated testing means that subjects in later trials are more experienced (Wagner, 1998; Reinhold & Schielzeth, 2015). This can influence choice behaviour in several ways. First, preferences may be influenced by the distribution of trait options previously encountered (e.g. Bakker & Milinski, 1991; Bateman & Fleming, 2006). For example, in the guppy (*Poecilia reticulata*), males prefer large females in sequential choice tests only after encountering females of variable size (Jordan & Brooks, 2012). If mating does not reduce receptivity, already-mated individuals can also attempt to 'trade-up' by becoming choosier in later trials, at little cost to their reproductive success (e.g. Pitcher *et al*., 2003). Prior experience can be controlled for either by fully randomising the order that options are presented in, or by ensuring that all subjects are presented with the same stimuli (Martin & Bateson, 2007). Second, subjects may become progressively more or less choosy with each successive mating opportunity, due to changes in the perceived costs and benefits of choice (Section II.2). For example, encountering multiple mates could lead to a higher perceived mate encounter rate, which could make subjects choosier.

VI. NUMBER OF OPTIONS DURING CHOICE TESTS

I have already considered the difference between being presented with a single option or multiple options (Section IV). Next I ask: when multiple options are presented during choice tests, how does the number of options a chooser is presented with influence choice? The most common simultaneous choice design is a two-choice design, although some studies, especially those considering acoustic communication, may present subjects with three or more options simultaneously (e.g. Beckers & Wagner, 2011; Forstmeier & Birkhead, 2004; Hall, Lindholm & Brooks, 2004; Brooks & Endler, 2001). Such multiple-option designs may bring several advantages over simpler no-choice or two-choice tests, but can introduce their own biases and problems.

(1) Advantages of using multiple options

There are three main advantages to presenting subjects with more than two options. The first is practical: when testing for relative preferences, fewer choice trials can be performed overall if each subject is presented with multiple options (Raffa *et al*., 2002). This can reduce the time needed to perform an experiment, or reduce the need for subjects to be tested more than once (although multiple no-choice tests are better for measuring absolute preferences; Section IV.1). Second, the more options used the greater our ability to reconstruct complex preference shapes (Section V.1). Both of these advantages only apply to tests that record the relative or rank differences between stimuli, by using a continuous measure of preference. This is because dichotomous choice outcomes only provide information on which stimulus is most preferred: a relative, continuous measure of preference is needed to 'fill in' the attractiveness of less-preferred options.

The third advantage of using a multiple-choice design is that they may often be more ecologically realistic than two-choice designs, by more closely reflecting the demographic conditions of natural populations. This is undoubtedly the case for species in which high-density male choruses are common (Beckers & Wagner, 2011). However, estimating the number of mates that are typically encountered in wild populations is difficult and is known only for well-studied species. A survey of 20 studies that recorded the number of males sampled by females in wild or free-ranging populations found an average of just 4.5 males sampled per female, and for all but three species the average number of males sampled was less than 10 (Roff & Fairbairn, 2014). This included three amphibian species with chorusing males, which all had low amounts of mate sampling (average of 1.3–2.6 males assessed per female). However, in these studies sampling was defined as close association or physical contact with a mate (e.g. Arak, 1988; Meuche *et al*., 2013). This excludes long-range assessment which could influence choice, and potentially underestimates the number of males that are perceived by sampling females. However, even in chorusing species it is not guaranteed that females are able to assess mates simultaneously. For example, in crickets the architecture of the auditory neurons leads to selective filtering out of quieter sounds played alongside louder ones (Pollack, 1988; Römer & Krusch, 2000), or those immediately following a stimulus (Greenfield, 2005, 2015; Section VII.2), presumably to reduce the cognitive load associated with trying to assess or locate multiple signals (Nityananda, 2016). Further, even though acoustic signals may travel far, male spacing of non-chorusing species in the wild may be such that females rarely hear multiple males simultaneously (e.g. Deb & Balakrishnan, 2014). Nevertheless, although admittedly based on a small sample, this survey suggests that, from an ecological perspective, two-choice tests clearly represent unnaturally low mate sampling for many species. Encouragingly, using only 3–5 options may be enough to mitigate this problem.

(2) Differences in the strength of choice

The number of options available during choice tests could influence the measurement of choice in several ways. First, the statistical power to detect a significant preference may be reduced as more options are added, especially when preferences are weak, the available options do not differ greatly in attractiveness, or assessment is prone to error. This means that larger sample sizes are needed to detect a given effect (Raffa *et al.*, 2002). Second, the number of options could influence the behaviour of subjects. For example, by varying the number of options, experiments simultaneously increase the perceived population density, and make the sex ratio biased towards the non-chooser sex (Section II.2). Both factors have the potential to increase the strength of mating preferences, by influencing the perceived mate encounter rate of subjects (Jennions & Petrie, 1997; Ah-King & Gowaty, 2016). It is well supported that the operational sex ratio significantly influences the strength of sexual selection (Janicke & Morrow, 2018). However, because most studies of sexual selection measure mating outcomes, it is not clear if this effect is driven by changes in mate choice or in the strength of intrasexual competition between potential mates (Hayes *et al*., 2016). Studies that compare mating preferences between choice tests with varying numbers of options, while also preventing intrasexual competition, are rare. In the broadnosed pipefish (*Syngnathus typhle*), males show stronger preferences when more females are available during choice tests (Berglund, 1994, 1995); however this pattern is not seen in two species of fiddler crabs (Hayes *et al*., 2016; Kawano & Henmi, 2016).

Third, increasing the number of options may impair the ability of subjects to assess and compare them. Studies of acoustic communication in animals frequently suggest that presenting more options simultaneously reduces the accuracy of choice, that is, the ability to identify the best option successfully (Hutchinson, 2005; Ryan *et al*., 2009). For example, in the painted reed frog (*Hyperolius marmoratus*), the ability of a female to locate the loudest speaker was reduced when four speakers were broadcasting compared to two (Bishop, Jennions & Passmore, 1995). This effect could arise because acoustic signals can interfere with each other (Forrest, 1994; Schwartz, Buchanan & Gerhardt, 2001; Greenfield, 2015). This is supported by the observation that female mating preferences are typically weaker when an extreme form of acoustic interference, white noise, is broadcast over calling males (e.g. Swaddle & Page, 2007; Bee & Schwartz, 2009).

It is less clear how the number of choice options influences choice for non-acoustic stimuli. This distinction is important, because visual or olfactory signals generally do not interfere with each other to the extent that acoustic signals do (but see Section VII.2). However, there are several other cognitive or perceptual issues which are exacerbated when multiple options are presented to subjects. For example, adding more options may reduce the difference in trait value between each option, making choice more difficult (Section VII.3). Animals may also take more time to choose when presented with multiple options (Bateson & Healy, 2005; Ryan *et al*., 2009). Given a universal speed–accuracy trade-off in decision making (Chittka, Skorupski & Raine, 2009; Gomez *et al*., 2011), assessment of multiple options may therefore impact decision accuracy when time is limited. Studies have suggested that humans find it harder to choose when simultaneously presented with a large number of options, leading to lower choice satisfaction and motivation to choose (known as the ‘choice overload’ or ‘overchoice’ effect: Schwartz, 2004; Hutchinson, 2005). This effect has been seen when choosing dating partners (Lenton, Fasolo & Todd, 2009; Lenton & Francesconi, 2011) and food options (Iyengar & Lepper, 2010), and is assumed to arise due to the cognitive difficulties associated with comparing many options at once, rather than interference (Hutchinson, 2005). However, a meta-analysis of the consumer choice literature suggests that across studies the average effect size is close to zero (Scheibehenne, Greifeneder & Todd, 2010), and it is unclear whether such effects are relevant to animal decision-making. Importantly, this cognitive impairment is not predicted for species that exhibit sequential choice, although there may be other cognitive difficulties associated with comparing options to previously encountered ones from memory (e.g. a best-of-n strategy: Janetos, 1980; Ryan *et al*., 2009) that could apply here. In Section VII.3 I discuss a range of cognitive issues that animals may encounter when comparing multiple options at once, which arise due to differences in the size, shape and location of options (rather than the number of options *per se*).

VII. STIMULI PRESENTED DURING TESTS

The final design consideration relates to the stimuli subjects are presented with during mate choice tests, with a specific focus on how these stimuli are perceived and assessed by subjects.

(1) Stimulus types

Stimulus types fall into four main categories: (1) live animals, (2) live animals with some phenotypic manipulation, (3) natural stimuli separated from the animals that produced them (e.g. calls and chemical cues), and (4) fully artificial stimuli (e.g. models, videos, synthesised acoustic or chemical cues). A less prevalent approach is to modify the environment to alter the perception of signals, for example by changing the available light spectrum or masking a frequency band in acoustic trials (Rosenthal, 2017). I will not consider this approach further, although I note that it is especially useful in order to confirm experimentally that a given trait affects mate choice (e.g. Milinski & Bakker, 1990; Kodric-Brown & Johnson, 2002; Kingston, Rosenthal & Ryan, 2003).

The advantage of using live animals as stimuli is ecological realism; this approach gets us closest to how the sexes interact in natural populations. However, this realism is offset by several potential disadvantages, the main problem being the uncontrolled variation associated with animal subjects. For example, stimulus animals may differ in physiological condition or receptivity in ways that are hard to detect, but might affect chooser responses. Further, stimulus animals may interact with choosers, even when direct contact is not possible, in ways that influence choices (Martin & Bateson, 2007). One way to exclude this possibility is the use of one-way screens, so that stimulus animals cannot see choosers (e.g. Forsgren, 1992; Hunt *et al*., 1997; Bisazza, Vaccari & Pilastro, 2001). Additionally, the use of live animal stimuli may be unsuitable if the aim is to determine exactly which traits or signals are being chosen, because of the difficulty of determining causality (Section III.1), especially when signalling is multimodal (Section VII.2). However, this problem can be mitigated to some extent by experimental manipulation of live animals, either by the modification (e.g. Andersson, 1982; Hebets, 2003; Conrad & Ayasse, 2015) or complete removal (e.g. Petrie & Halliday, 1994; Dakin & Montgomerie, 2013) of existing traits, or the addition of novel traits (e.g. Basolo, 1990*b*; Burley & Symanski, 1998). This is a powerful technique as long as the manipulation can be performed without altering other aspects of the stimulus animal's behaviour, or the appearance of other signals (such as the relative size of a trait: Gerhardt, 1992).

In many cases, animal signals can be isolated from the animals themselves during mate choice experiments, facilitated by the use of proxy behavioural measures of preference which do not require that mates directly interact (Section III.2). For example, speakers can be used to broadcast acoustic signals (known as playback experiments), and chemical cues can be added to the environment or placed on a substrate. This has the advantage that responses can be tested to a focal trait in isolation, and allows us to establish a causal link between the two. Video playback is a powerful technique which allows experimenters to test subject responses to complex, moving visual scenes (D'Eath, 1998). Experimenters need to be sure that subjects respond to such isolated stimuli, as this form of presentation might not work in species with complex courtship displays that involve active interaction and feedback with mates (although modern, interactive playback may partly solve this problem: King, 2015). Nevertheless, a large body of research suggests that subjects typically respond to these stimuli as if they were faced with real animals (Powell & Rosenthal, 2016). As with proxy measures of preferences (Section III.2), the responses of subjects to these stimuli should be validated by comparing them to actual choice outcomes, if we are interested in the fitness consequences of behaviour (see also Section VII.3).

Researchers have the most control over stimuli that are artificially created. Examples include dummies, synthetic calls and chemical cues (Powell & Rosenthal, 2016), and animation and virtual reality (Gierszewski *et al*., 2017; Chouinard-Thuly et al, 2017). This approach allows for the creation of stimuli that are simplified compared to natural stimuli, or which differ along only a single axis of variation; making them suitable for studies interesting in identifying which aspects of a sexual signal are salient during choice. Artificial stimuli can also be created which vary several traits simultaneously (e.g. Brooks *et al*., 2005; Mautz *et al*., 2013); such an approach is potentially powerful as it allows experimenters to decouple natural correlations between traits which might obscure preferences, and to test which specific combinations of trait values are most attractive. Conversely, researchers can create stimuli which lie outside the range of trait values typically encountered in the wild (Powell & Rosenthal, 2016), and to explore whether preferences are open-ended, or reflect ancestral sensory biases (Basolo, 1990*b*; Ryan & Keddy-Hector, 1992). Testing preferences for stimuli slightly above and below the natural range can help to describe more fully how choosers impose selection on chosen traits, even if other factors limit the ultimate expression of those traits (Kilmer *et al*., 2017). Importantly, isolated signals can also be used to remove within-individual variability in stimuli (Section V.1), for example by testing all subjects with the same recorded or synthetic call, so that preferences may be easier to detect (Powell & Rosenthal, 2016). However, testing subjects with one or a small number of stimuli is a form of pseudoreplication that reduces the ability to generalise the results of any given experiment (McGregor *et al*., 1992; Kroodsma *et al*., 2001). Studies should use either multiple exemplars in the case of natural stimuli, or multiple synthetic stimuli with at least some random variation introduced, to reduce such pseudoreplication (McGregor *et al*., 1992). Finally, and perhaps most importantly, using artificial stimuli introduces the associated risk that subjects do not perceive the stimulus as a sexual signal, or even as a signal at all (Section VII.3). This could lead to reduced overall responses of subjects when compared to natural stimuli, or even active aversion (Martin & Bateson, 2007).

(2) Signal modalities

Another important consideration is the modality of the signals that subjects are assessing, as different modalities are prone to different methodological constraints. For example, modern audio-editing software allows for the easy modification of acoustic signals, or the creation of synthetic ones that are indistinguishable from unmodified or natural signals (both to humans and animal subjects: Gerhardt, 1978, 1981). This is much more difficult for all but the simplest olfactory or visual signals (Rosenthal & Ryan, 2000). Additionally, the signals used in acoustic and chemical courtship are often highly context specific, so that they are only ever produced during mating interactions (Rosenthal, 2017), and animals may have receptors that are incredibly finely tuned to the sexual signals produced by mates (e.g. Leinders-Zufall *et al*., 2000). Subjects may therefore be highly sensitive to changes in acoustic and chemical signals, compared to variation in visual signals such as body size which are not context specific, and indeed may not be actively ‘broadcast’ by potential partners. However, the extent to which this affects mate choice remains unclear. Visual stimuli require several additional methodological considerations, such as the need to standardise the colour and brightness of lighting and the background used during tests (Lynn & Cole, 2019).

Signal modalities also differ in the extent to which interference is likely. Interference is common for acoustic stimuli, most notably in relation to volume: louder signals obscure quieter ones, an effect which arises due to both the physics and neurology of sound detection (Section VI.2). Acoustic signals are also more prone to degradation than visual signals over medium to long distances, which can be a problem for playback experiments in the field (Gerhardt, 1992; Rosenthal & Ryan, 2000). Both acoustic and visual signals are also vulnerable to temporal interference. For example, group-signalling species often exhibit synchronicity in the production of rhythmic, repetitive signals (Greenfield, 2005). Small differences in call timing generate ‘leaders’ and ‘followers’ in relation to each signal pulse, and females prefer leading signals in acoustically signalling anurans and orthopterans (Greenfield, 2005), and visually signalling fireflies and fiddler crabs (Lewis & Cratsley, 2008; Reaney *et al*., 2008), an effect which may be driven by neural constraints (Greenfield, 2005, 2015).

Mate choice experiments typically test for mating preferences for a single trait at a time. However, sexual signalling may often be multimodal (Candolin, 2003; Mitoyen, Quigley & Fusani, 2019). Studies comparing choice behaviour between subjects presented with unimodal or multimodal displays have shown that the presentation of multimodal signals may increase the overall response rate (Bailey, 2011; Uetz, Roberts & Taylor, 2009; Reichert & Höbel, 2015), improve choice accuracy when options are similar (Gomez *et al*., 2011), and lead to weaker mating preferences for any single component (Reichert & Höbel, 2015). Given that multimodal signals represent a more ecologically realistic scenario, this latter finding raises the possibility that laboratory experiments consistently overestimate the strength of sexual selection on any given display component when it is tested in isolation, compared to the natural, multimodal situation. However, despite the ubiquity of multimodal signalling, relatively few studies to date have tested for these effects (Mitoyen *et al*., 2019).

(3) Perceptual considerations

Mate choice experiments assume that subjects are able effectively to perceive and distinguish between the options they are presented with. However, certain types (or combinations) of stimuli may be more easily perceived than others (Bateson & Healy, 2005; Ryan & Cummings, 2013; Akre & Johnsen, 2014; Kelley & Kelley, 2014), so that a lack of mate choice could arise due to perceptual or cognitive constraints (Section VI.2). Assumptions about perception are especially problematic with artificial stimuli, and we need to ensure that animals perceive what we think they are perceiving. Animal perception can be surprisingly different to our own: colour vision, for example, is highly variable across species (e.g. Stoddard *et al*., 2018). Video images are especially vulnerable to this problem because video screens are designed for human viewers: animals may differ, for example, in how they perceive colour, image depth or screen refresh rate (see D’Eath, 1998; Oliveira *et al*., 2000; Powell & Rosenthal, 2016). Ideally, some form of validation is needed to confirm that artificial stimuli are as attractive as natural stimuli (Lahti, 2015, e.g. Gerhardt, 1978, 1981; Moravec, Striedter & Burley, 2010; Gierszewski *et al*., 2017). Artificial stimuli can be validated by comparison of subjects’ responses to natural stimuli, or mechanistic knowledge of animal perception or cognition (e.g. modelling animal colour vision using the presence of receptors: Stoddard *et al*., 2018), and validation becomes more important the more stimuli differ from their natural models (Lahti, 2015).

Even for natural stimuli, there are general perceptual or cognitive processes that can influence assessment and choice, and are relevant for the design of mate choice experiments. An important perceptual issue during simultaneous assessment relates to the ability to discriminate between potential choice options: presenting subjects with options that are too similar may reduce their ability to discriminate (Hutchinson, 2005; Raffa *et al*., 2002), especially when animals are time-stressed (Sullivan, 1994; Schwartz *et al*., 2001; Chittka *et al*., 2009). For example, female tree frogs are less likely to identify the highest quality signal, and take longer to choose, when signals are more similar (Gomez *et al*., 2011). The difference between options is unavoidably reduced when more options are included in a test and the absolute range remains the same. Tests should also give subjects sufficient time to assess and compare options before being asked to make a choice (Schwartz, Huth & Hutchin, 2004; Chittka *et al*., 2009; Section VI.2), and in cases where the time needed for assessment is unknown, experimenters should err on the side of caution by giving subjects plenty of time.

The ability to distinguish between options may also vary depending on the magnitude of stimuli, because animals typically compare stimuli based on proportional differences in stimulus magnitude (Ryan & Cummings, 2013; Akre & Johnsen, 2014). When this is the case, discrimination is more difficult between options of large magnitude, because a larger absolute difference is needed in order to perceive the same proportional change. This fundamental property of psychophysics is known as Weber’s law, and applies to animals across a broad taxonomic range, from humans to bees, and across a range of sensory modalities (Akre & Johnsen, 2014). This effect may also be relevant to mate choice: despite the fact that animals often have open-ended preferences for greater intensity signals (Ryan & Keddy-Hector, 1992), the ability to distinguish between options should be reduced as signal intensity increases. For example, Túngara frog (*Physalaemus pustulosus*) females typically prefer more elaborate male calls, but experimental addition of call components increases male attractiveness to a lesser degree as call elaboration increases, an effect consistent with Weber’s law (Akre *et al*., 2011).

Visual perception in humans is vulnerable to visual illusions, in which the appearance of the background or other options can influence how the size, shape, colour or movement of objects are perceived (Kelley & Kelley, 2014; Lynn & Cole, 2019). If animals are similarly vulnerable, then visual illusions have the potential to influence mate assessment, especially when assessment is comparative (in the case of sequential assessment any cognitive issues are a problem of memory rather than illusion). For example, humans are vulnerable to the Ebbinghaus–Titchener illusion, where shapes appear larger or smaller depending on the size of surrounding shapes (Bateson & Healy, 2005; Ludwig & Pollet, 2014; Kelley & Kelley, 2014). There is some evidence that this illusion influences mate assessment in animals: the same average-sized male banana fiddler crab (*Uca mjoebergi*) attracts more females when signalling directly next to small males compared to when signalling next to large ones, even when the size distribution of available mates is the same (Callander *et al*., 2012). Indeed, in some cases animals have been shown to perceive visual illusions in a similar way to humans (Murayama *et al*., 2012), although other studies find no effect, or an effect in the opposite direction (e.g. Parron & Fagot, 2007; Lucon-Xiccato *et al*., 2019). This is, of course, intrinsically informative if the purpose of the study is to examine animal cognition, but it illustrates the risks associated with assuming that animal perception is similar to our own.

VIII. RECOMMENDATIONS

I have highlighted how several fundamental aspects of experimental design may influence the measurement of mate choice, either by influencing chooser behaviour or our ability to measure that behaviour. These effects should be taken into account when designing experiments, or comparing empirical results across studies. Nevertheless, variation in the design of mate choice experiments is unavoidable, given the wide range of species and questions being tested. I therefore emphasise that there is no ‘correct’ way to test for mate choice; experimental design is always an exercise in compromise. Instead, I have several general recommendations for future work. First, for a variety of reasons (e.g. weak overall preferences, complex preference functions, low repeatability of preferences both within and among individuals), the measurement of mate choice may be particularly difficult when sample sizes are small. Therefore statistical power should be an important consideration for mate choice experiments if we are to obtain robust results. More studies should also attempt to examine multimodal and multivariate mating preferences, given that such preferences are likely widespread and ecologically relevant. Reporting of experimental methods also needs to be improved (McGregor *et al*., 1992; Kilkenny *et al*., 2010; Griffith *et al*., 2017), given how sensitive mate choice behaviour may be to the environment experienced by animals before and during choice tests.

If the aim of a study is to record mating preferences under ‘ideal’ conditions, then in general, a preference function approach, combined with estimates of within-individual repeatability, is the best approach to characterise mating preferences fully (Wagner, 1998; Rosenthal, 2017), and more studies should attempt this if possible. The exact experimental setup best suited to reconstructing a preference function depends on three factors: the complexity of the preference shape you are trying to reconstruct, the level of preference you are concerned with (individual or population level) and the choice design used (Table 4). For simple, linear preferences, a single, two-choice test will generally suffice for both individual- and population-level preferences. For complex preferences, the number of choice trials needed per individual depends on the number of options presented during the test: the more options per test, the fewer trials needed. The number of no-choice trials needed per individual depends on the preference level. Only one trial per individual is required for population-level preferences (as long as some individuals are tested with different stimuli), regardless of preference shape. However, for individual-level preferences, the number of no-choice trials per individual depends on the complexity of the preference: the more complex the preference, the more trials are needed.

Alternatively, if the aim is to estimate how choice occurs in natural populations, ecological and biological relevance should often trump other considerations when planning experiments, even if this leads to known experimental issues, or weaker mating preferences compared to more ‘ideal’ conditions. It is unavoidable that experimental design is limited by the biology of the organism being examined: species vary greatly in what sexual signals they transmit, what behaviours they exhibit, what stimuli they respond to, and the sampling strategies and cognitive rules they use to compare mates. This variability is reflected in the experimental designs used to test for mate choice, and may unavoidably lead to some experimental setups being group or species specific. Importantly, if the aim of an experiment is to infer something about the ecology or evolution of a species in the wild, then the experimental conditions should aim to match the conditions under which mates are encountered and assessed in natural populations. For most species we still have very little information on how this occurs. However, without this knowledge laboratory studies run the risk of over- or under-estimating the strength of mating preferences, if they employ an experimental approach that is ecologically unrealistic. Importantly, there is rarely one ‘true’ ecologically relevant context for any given species, due to spatial or temporal variability. For example, mate encounter may often be stochastic, so that mates are sometimes encountered both sequentially and simultaneously. If there are reasons to believe that both types of mate encounter are reasonably common, then the results of both sequential and simultaneous tests will be informative.

These two approaches ('ideal' *versus* 'ecologically relevant') represent extremes which may be in conflict: ecological relevance is often the antithesis to ideal, controlled conditions. However, neither approach is superior; they just help us answer different questions. Because of this, an integrative approach, combining results from multiple types of test, will result in a more complete and robust characterisation of mate choice for any given species (Powell & Rosenthal, 2016). Multiple experiments, which each vary a small number of 'confounding' factors, can be directly compared, in order to quantify how each factor influences the expression of choice behaviour. This applies, for example, to the dichotomy between proxy measures of preference and mating outcomes; comparing results using these two approaches in the same species should be informative, as each brings benefits that the other lacks (Rosenthal, 2017).

Finally, we need more quantitative estimates of how different experimental setups influence mate choice, either within or across species. This has been done in only a few cases to my knowledge (Dougherty & Shuker, 2015*b*; Nieberding & Holveck, 2017; Griffith *et al*., 2017). Consequently, we lack reliable estimates of how choice is influenced by the type of behavioural measure recorded (proxy *versus* mating outcome), the number of options presented, or the type of stimuli (stimulus type or signal modality). It may be that these different designs do not influence choice behaviour in any consistent way, or they may systematically lead to stronger preferences in certain designs. Meta-analysis in particular is useful for testing hypotheses when there are many published studies that suffer individually from low statistical power. Although broad taxonomic analyses will be most generalisable, those focused on single species or groups have greater resolution and are also valuable. These estimates should allow us more effectively to weigh up different design considerations when planning our experiments and, when practical considerations force our hand, to estimate to what extent our measurements would change under different conditions. Without this information, we are unable to determine to what extent experimental conditions are responsible for the variability in mate choice and mating behaviour observed both within and across species.

This lack of quantification has two main consequences. First, our comparison of mate choice experiments is hindered until we can determine effectively to what extent experimental design may confound our measurements of choice. For a research field that relies on results from a diverse taxonomic range, this is a problem. Second, it hinders our understanding of how mate choice acts in natural populations. For example, our laboratory experiments could consistently under- or overestimate the strength of mating preferences in wild populations, if we routinely use experimental setups that are entirely unrealistic. This is a problem if we want to understand mate choice as a key component of sexual selection, and a driver of trait evolution in both the choosing and chosen sex (Rosenthal, 2017). Further, given the importance of mate choice for initiating and maintaining reproductive isolation between animal populations, this lack of knowledge also hinders our understanding of speciation (Andersson, 1994; Ritchie, 2007). Improved reporting standards, and a greater consideration of the importance of experimental design when designing mate choice experiments, will help us begin to solve these problems.

IX. CONCLUSIONS

(1) The mating choices animals make are inextricably linked to their state, their prior experience, and the environments they are tested in. Small changes in experimental conditions can potentially lead to large differences in the expression of mate choice, and experimenters should try to control for these effects when designing experiments.

(2) There are many ways we attempt to measure mate choice, but not all of them reflect what we intend to measure equally well. Importantly, different experimental approaches may influence how animals behave, what aspects of mating behaviour we are measuring, or our ability to detect statistically significant differences in behaviour.

(3) We need to consider carefully what animals are perceiving during our experiments, both in terms of the signals we present to them, or any unintended cues which could alter their decisions.

(4) In most cases it is still unclear how different experimental approaches quantitatively influence the expression or measurement of animal mate choice. Formal testing of these effects will allow us to weigh up different design considerations more effectively when planning our experiments and estimate to what extent our measurements would change under different conditions.

(5) Consideration of these effects is important if we are to compare mate choice behaviour across species and experimental designs, or if we want to understand the evolution and expression of mate choice in natural environments.

X. ACKNOWLEDGEMENTS

I would like to thank Gil Rosenthal, William Wagner, Michael Jennions and Zenobia Lewis for comments which greatly improved this review. This work was supported by the Leverhulme Trust (ECF-2018-427) and the Association for the Study of Animal Behaviour.

XI. REFERENCES

AH- KING, M. & GOWATY, P. A. (2016). A conceptual review of mate choice: stochastic demography, within-sex phenotypic plasticity, and individual flexibility. *Ecology and Evolution* **6**, 4607–4642.

AHARON, I., ETCOFF, N., ARIELY, D., CHABRIS, C. F., O'CONNOR, E. & BREITER, H. C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron* **32**, 537–551.

AKRE, K. L., FARRIS, H. E., LEA, A. M., PAGE, R. A. & RYAN, M. J. (2011). Signal perception in frogs and bats and the evolution of mating signals. *Science* **333**, 751–752.

AKRE, K. L. & JOHNSEN, S. (2014). Psychophysics and the evolution of behavior. *Trends in Ecology & Evolution* **29**, 291–300.

ALLISON, J. D. & CARDÉ, R. T. (2008). Male pheromone blend preference function measured in choice and no-choice wind tunnel trials with almond moths, *Cadra cautella*. *Animal Behaviour* **75**, 259–266.

AMY, M., SALVIN, P., NAGUIB, M. & LEBOUCHER, G. (2015). Female signalling to male song in the domestic canary, *Serinus canaria*. *Royal Society Open Science* **2**: 140196.

ANDERSSON, M. (1982). Female choice selects for extreme tail length in a widowbird. *Nature* **299**, 818–820.

ANDERSSON, M. B. (1994). *Sexual selection*. Princeton University Press, Princeton.

ARAK, A. (1988). Female mate selection in the natterjack toad: active choice or passive atraction? *Behavioral Ecology and Sociobiology* **22**, 317–327.

ARNQVIST, G. & ROWE, L. (2005). *Sexual conflict*. Princeton University Press, Princeton.

BAILEY, N. W. (2008). Love will tear you apart: different components of female choice exert contrasting selection pressures on male field crickets. *Behavioral Ecology* **19**, 960–966.

BAILEY, N. W. (2011). Mate choice plasticity in the field cricket *Teleogryllus oceanicus*: effects of social experience in multiple modalities. *Behavioral Ecology and Sociobiology* **65**, 2269–2278.

BAILEY, N. W. & ZUK, M. (2008). Acoustic experience shapes female mate choice in field crickets. *Proceedings of the Royal Society B: Biological Sciences* **275**, 2645–2650.

BAKKER, T. C. & MILINSKI, M. (1991). Sequential female choice and the previous male effect in sticklebacks. *Behavioral Ecology and Sociobiology* **29**, 205–210.

BARRY, K. L., HOLWELL, G. I. & HERBERSTEIN, M. E. (2010). Multimodal mate assessment by male praying mantids in a sexually cannibalistic mating system. *Animal Behaviour* **79**, 1165–1172.

BARRY, K. L. & KOKKO, H. (2010). Male mate choice: why sequential choice can make its evolution difficult. *Animal Behaviour* **80**, 163–169.

BASOLO, A. L. (1990*a*). Female preference for male sword length in the green swordtail, *Xiphophorus helleri* (Pisces: Poeciliidae). *Animal Behaviour* **40**, 332–338.

BASOLO, A. L. (1990*b*). Female preference predates the evolution of the sword in swordtail fish. *Science* **250**, 808–810.

BATEMAN, P. W. & FLEMING, P. A. (2006). Males are selective too: mating, but not courtship, with sequential females influences choosiness in male field crickets (*Gryllus bimaculatus*). *Behavioral Ecology and Sociobiology* **59**, 577–581.

BATESON, M. & HEALY, S. D. (2005). Comparative evaluation and its implications for mate choice. *Trends in Ecology & Evolution* **20**, 659–664.

BAXTER, C., MENTLIK, J., SHAMS, I. & DUKAS, R. (2018). Mating success in fruit flies: courtship interference versus female choice. *Animal Behaviour* **138**, 101–108.

BEACH, F. A. (1976). Sexual attractivity, proceptivity, and receptivity in female mammals. *Hormones and Behavior* **7**, 105–138.

BEATTY, C. & FRANKS, D. (2012). Discriminative predation: simultaneous and sequential encounter experiments. *Current Zoology* **58**, 649–657.

BECKERS, O. M. & WAGNER, W. E., JR. (2011). Mate sampling strategy in a field cricket: evidence for a fixed threshold strategy with last chance option. *Animal Behaviour* **81**, 519–527.

BEE, M. A. & SCHWARTZ, J. J. (2009). Behavioral measures of signal recognition thresholds in frogs in the presence and absence of chorus-shaped noise. *The Journal of the Acoustical Society of America* **126**, 2788–2801.

BELL, A. M., HANKISON, S. J. & LASKOWSKI, K. L. (2009). The repeatability of behaviour: a meta-analysis. *Animal Behaviour* **77**, 771–783.

BERGLUND, A. (1994). The operational sex ratio influences choosiness in a pipefish. *Behavioral Ecology* **5**, 254–258.

BERGLUND, A. (1995). Many mates make male pipefish choosy. *Behaviour* **132**, 213–218.

BISAZZA, A., VACCARI, G. & PILASTRO, A. (2001). Female mate choice in a mating system dominated by male sexual coercion. *Behavioral Ecology* **12**, 59–64.

BISHOP, P. J., JENNIONS, M. D. & PASSMORE, N. I. (1995). Chorus size and call intensity: female choice in the painted reed frog, *Hyperolius marmoratus*. *Behaviour* **132**, 721–731.

BONDURIANSKY, R. (2001). The evolution of male mate choice in insects: a synthesis of ideas and evidence. *Biological Reviews* **76**, 305–339.

BOOKSMYTHE, I., DETTO, T. & BACKWELL, P. R. (2008). Female fiddler crabs settle for less: the travel costs of mate choice. *Animal Behaviour* **76**, 1775–1781.

BOOKSMYTHE, I., JENNIONS, M. D. & BACKWELL, P. R. (2011). Male fiddler crabs prefer conspecific females during simultaneous, but not sequential, mate choice. *Animal Behaviour* **81**, 775–778.

BROOKS, R. & ENDLER, J. A. (2001). Female guppies agree to differ: phenotypic and genetic variation in mate- choice behavior and the consequences for sexual selection. *Evolution* **55**, 1644–1655.

BROOKS, R., HUNT, J., BLOWS, M. W., SMITH, M. J., BUSSIÈRE, L. F. & JENNIONS, M. D. (2005). Experimental evidence for multivariate stabilizing sexual selection. *Evolution* **59**, 871–880.

BURLEY, N. T. & SYMANSKI, R. (1998). “A taste for the beautiful”: Latent aesthetic mate preferences for white crests in two species of Australian grassfinches. *The American Naturalist* **152**, 792–802.

BUSH, S. L., GERHARDT, H. C. & SCHUL, J. (2002). Pattern recognition and call preferences in treefrogs (Anura: Hylidae): a quantitative analysis using a no-choice paradigm. *Animal Behaviour* **63**, 7–14.

BYERS, J. A., WISEMAN, P. A., JONES, L. & ROFFE, T. J. (2005). A large cost of female mate sampling in pronghorn. *The American Naturalist* **166**, 661–668.

CALLANDER, S., HAYES, C. L., JENNIONS, M. D. & BACKWELL, P. R. (2012). Experimental evidence that immediate neighbors affect male attractiveness. *Behavioral Ecology* **24**, 730–733.

CANDOLIN, U. (2003). The use of multiple cues in mate choice. *Biological Reviews* **78**, 575–595.

CANDOLIN, U., SALESTO, T. & EVERS, M. (2007). Changed environmental conditions weaken sexual selection in sticklebacks. *Journal of Evolutionary Biology* **20**, 233–239.

CHENOWETH, S. F. & BLOWS, M. W. (2005). Contrasting mutual sexual selection on homologous signal traits in *Drosophila serrata*. *The American Naturalist* **165**, 281–289.

CHITTKA, L., SKORUPSKI, P. & RAINE, N. E. (2009). Speed–accuracy tradeoffs in animal decision making. *Trends in Ecology & Evolution* **24**, 400–407.

CHIVERS, M. L., SETO, M. C., LALUMIÈRE, M. L., LAAN, E. & GRIMBOS, T. (2010). Agreement of self-reported and genital measures of sexual arousal in men and women: A meta-analysis. *Archives of Sexual Behavior* **39**, 5–56.

CHOUINARD-THULY, L., GIERSZEWSKI, S., ROSENTHAL, G. G., READER, S. M., RIEUCAU, G., WOO, K. L., GERLAI, R., TEDORE, C., INGLEY, S. J. & STOWERS, J. R. (2017). Technical and conceptual considerations for using animated stimuli in studies of animal behavior. *Current Zoology* **63**, 5–19.

CONRAD, T. & AYASSE, M. (2015). The role of vibrations in population divergence in the red mason bee, *Osmia bicornis*. *Current Biology* **25**, 2819–2822.

COTTON, S., SMALL, J. & POMIANKOWSKI, A. (2006). Sexual selection and condition-dependent mate preferences. *Current Biology* **16**, R755–R765.

COULDRIDGE, V. & ALEXANDER, G. (2001). Does the time spent near a male predict female mate choice in a Malawian cichlid? *Journal of Fish Biology* **59**, 667–672.

COYNE, J. A., ELWYN, S. & ROLÁN-ALVAREZ, E. (2005). Impact of experimental design on Drosophila sexual isolation studies: direct effects and comparison to field hybridization data. *Evolution* **59**, 2588–2601.

CUMMINGS, M. & MOLLAGHAN, D. (2006). Repeatability and consistency of female preference behaviours in a northern swordtail, *Xiphophorus nigrensis*. *Animal Behaviour* **72**, 217–224.

CUMMINGS, M. E. (2015). The mate choice mind: studying mate preference, aversion and social cognition in the female poeciliid brain. *Animal Behaviour* **103**, 249–258.

DAKIN, R. & MONTGOMERIE, R. (2013). Eye for an eyespot: how iridescent plumage ocelli influence peacock mating success. *Behavioral Ecology* **24**, 1048–1057.

DARWIN, C. R. (1871). *The descent of man, and selection in relation to sex.* John Murray, London.

DAVID, M. & CEZILLY, F. (2011). Personality may confound common measures of mate-choice. *Plos One* **6**, e24778.

DE JONG, M. C. & SABELIS, M. W. (1991). Limits to runaway sexual selection: the wallflower paradox. *Journal of Evolutionary Biology* **4**, 637–655.

D'EATH, R. B. (1998). Can video images imitate real stimuli in animal behaviour experiments? *Biological Reviews* **73**, 267–292.

DEB, R. & BALAKRISHNAN, R. (2014). The opportunity for sampling: the ecological context of female mate choice. *Behavioral Ecology* **25**, 967–974.

DIXSON, B. J., GRIMSHAW, G. M., LINKLATER, W. L. & DIXSON, A. F. (2011). Eye-tracking of men's preferences for waist-to-hip ratio and breast size of women. *Archives of Sexual Behavior* **40**, 43–50.

DOUGHERTY, L. R. & SHUKER, D. M. (2014). Pre-copulatory sexual selection in the seed bug *Lygaeus equestris*: a comparison of choice and no-choice paradigms. *Animal Behaviour* **89**, 207–214.

DOUGHERTY, L. R. & SHUKER, D. M. (2015*a*). Mate choice and mating decisions: a response to comments on Dougherty & Shuker. *Behavioral Ecology* **26**, 326–327.

DOUGHERTY, L. R. & SHUKER, D. M. (2015*b*). The effect of experimental design on the measurement of mate choice: a meta-analysis. *Behavioral Ecology* **26**, 311–319.

DUGATKIN, L. A. (1992). Sexual selection and imitation: females copy the mate choice of others. *The American Naturalist* **139**, 1384–1389.

EDWARD, D. A. (2015). The description of mate choice. *Behavioral Ecology* **26**, 301–310.

EDWARD, D. A. & CHAPMAN, T. (2011). The evolution and significance of male mate choice. *Trends in Ecology and Evolution* **26**, 647–654.

FORREST, T. (1994). From sender to receiver: propagation and environmental effects on acoustic signals. *American Zoologist* **34**, 644–654.

FORSGREN, E. (1992). Predation risk affects mate choice in a gobiid fish. *The American Naturalist* **140**, 1041–1049.

FORSTMEIER, W. & BIRKHEAD, T. R. (2004). Repeatability of mate choice in the zebra finch: consistency within and between females. *Animal Behaviour* **68**, 1017–1028.

FOWLER-FINN, K. D. & RODRÍGUEZ, R. L. (2013). Repeatability of mate preference functions in Enchenopa treehoppers (Hemiptera: Membracidae). *Animal Behaviour* **85**, 493–499.

GABOR, C. (1999). Association patterns of sailfin mollies (*Poecilia latipinna*): alternative hypotheses. *Behavioral Ecology and Sociobiology* **46**, 333–340.

GABOR, C. R., KRENZ, J. D. & JAEGER, R. G. (2000). Female choice, male interference, and sperm precedence in the red-spotted newt. *Behavioral Ecology* **11**, 115–124.

GAVRILETS, S., ARNQVIST, G. & FRIBERG, U. (2001). The evolution of female mate choice by sexual conflict. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **268**, 531–539.

GERHARDT, H. C. (1978). Temperature coupling in the vocal communication system of the gray tree frog, *Hyla versicolor*. *Science* **199**, 992–994.

GERHARDT, H. C. (1981). Mating call recognition in the barking treefrog (*Hyla gratiosa*): responses to synthetic calls and comparisons with the green treefrog (*Hyla cinerea*). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology* **144**, 17–25.

GERHARDT, H. C. (1992). Conducting playback experiments and interpreting their results. In *Playback and studies of animal communication* (ed. P. K. McGregor), pp. 59–77. Springer, Boston.

GIERSZEWSKI, S., MÜLLER, K., SMIELIK, I., HÜTWOHL, J.-M., KUHNERT, K.-D. & WITTE, K. (2017). The virtual lover: variable and easily guided 3D fish animations as an innovative tool in mate-choice experiments with sailfin mollies-II. Validation. *Current Zoology* **63**, 65–74.

GILLINGHAM, M. A., RICHARDSON, D. S., LØVLIE, H., MOYNIHAN, A., WORLEY, K. & PIZZARI, T. (2008). Cryptic preference for MHC-dissimilar females in male red junglefowl, *Gallus gallus*. *Proceedings of the Royal Society B: Biological Sciences* **276**, 1083–1092.

GOMEZ, D., THÉRY, M., GAUTHIER, A.-L. & LENGAGNE, T. (2011). Costly help of audiovisual bimodality for female mate choice in a nocturnal anuran (*Hyla arborea*). *Behavioral Ecology* **22**, 889–898.

GONÇALVES, D. M. & OLIVEIRA, R. F. (2003). Time spent close to a sexual partner as a measure of female mate preference in a sex-role-reversed population of the blenny *Salaria pavo* (Risso) (Pisces: Blenniidae). *Acta Ethologica* **6**, 1–5.

GRANT, P. R. & GRANT, B. R. (2018). Role of sexual imprinting in assortative mating and premating isolation in Darwin's finches. *Proceedings of the National Academy of Sciences* **115**, E10879–E10887.

GREENFIELD, M. D. (2005). Mechanisms and evolution of communal sexual displays in arthropods and anurans. *Advances in the Study of Behavior* **35**, 1–62.

GREENFIELD, M. D. (2015). Signal interactions and interference in insect choruses: singing and listening in the social environment. *Journal of Comparative Physiology A* **201**, 143–154.

GREENWAY, E. V. G., DOUGHERTY, L. R. & SHUKER, D. M. (2015). Mating failure. *Current Biology* **25**, R534–R536.

GRIFFITH, S. C., CRINO, O. L., ANDREW, S. C., NOMANO, F. Y., ADKINS- REGAN, E., ALONSO-ALVAREZ, C., BAILEY, I. E., BITTNER, S. S., BOLTON, P. E. & BONER, W. (2017). Variation in reproductive success across captive populations: methodological differences, potential biases and opportunities. *Ethology* **123**, 1–29.

GUBERNICK, D. J. & ADDINGTON, R. L. (1994). The stability of female social and mating preferences in the monogamous California mouse, *Peromyscus californicus*. *Animal Behaviour* **47**, 559–567.

HALL, M., LINDHOLM, A. K. & BROOKS, R. (2004). Direct selection on male attractiveness and female preference fails to produce a response. *Bmc Evolutionary Biology* **4**, 1.

HALLIDAY, T. (1983). The study of mate choice. In *Mate choice* (ed. P. Bateson), pp. 3–32. Cambridge University Press, Cambridge.

HAYES, C., CALLANDER, S., BOOKSMYTHE, I., JENNIONS, M. & BACKWELL, P. (2016). Mate choice and the operational sex ratio: an experimental test with robotic crabs. *Journal of Evolutionary Biology* **29**, 1455–1461.

HEAD, M. L., JACOMB, F., VEGA-TREJO, R. & JENNIONS, M. D. (2015). Male mate choice and insemination success under simultaneous versus sequential choice conditions. *Animal Behaviour* **103**, 99–105.

HEAD, M. L., LINDHOLM, A. K. & BROOKS, R. (2008). Operational sex ratio and density do not affect directional selection on male sexual ornaments and behavior. *Evolution* **62**, 135–144.

HEBETS, E. A. (2003). Subadult experience influences adult mate choice in an arthropod: Exposed female wolf spiders prefer males of a familiar phenotype. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 13390–13395.

HINGLE, A., FOWLER, K. & POMIANKOWSKI, A. (2001). The effect of transient food stress on female mate preference in the stalk–eyed fly *Cyrtodiopsis dalmanni*. *Proceedings of the Royal Society B: Biological Sciences* **268**, 1239–1244.

HOARE, D., KRAUSE, J., PEUHKURI, N. & GODIN, J. G. (2000). Body size and shoaling in fish. *Journal of Fish Biology* **57**, 1351–1366.

HOLMAN, L., HEAD, M. L., LANFEAR, R. & JENNIONS, M. D. (2015). Evidence of experimental bias in the life sciences: why we need blind data recording. *Plos Biology* **13**, e1002190.

HOLVECK, M.-J., GAUTHIER, A.-L. & NIEBERDING, C. M. (2015). Dense, small and male-biased cages exacerbate male–male competition and reduce female choosiness in *Bicyclus anynana*. *Animal Behaviour* **104**, 229–245.

HOLVECK, M.-J. & RIEBEL, K. (2007). Preferred songs predict preferred males: consistency and repeatability of zebra finch females across three test contexts. *Animal Behaviour* **74**, 297–309.

HUGHES, N. K., KELLEY, J. L. & BANKS, P. B. (2012). Dangerous liaisons: the predation risks of receiving social signals. *Ecology Letters* **15**, 1326–1339.

HUNT, J., BROOKS, R. & JENNIONS, M. D. (2005). Female mate choice as a condition-dependent life- history trait. *The American Naturalist* **166**, 79–92.

HUNT, S., CUTHILL, I. C., SWADDLE, J. P. & BENNETT, A. T. (1997). Ultraviolet vision and band-colour preferences in female zebra finches, *Taeniopygia guttata*. *Animal Behaviour* **54**, 1383–1392.

HUTCHINSON, J. (2005). Is more choice always desirable? Evidence and arguments from leks, food selection, and environmental enrichment. *Biological Reviews* **80**, 73–92.

IMMELMANN, K. (1975). Ecological significance of imprinting and early learning. *Annual Review of Ecology and Systematics* **6**, 15–37.

IYENGAR, S. S. & LEPPER, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology* **79**, 995–1006.

JANETOS, A. C. (1980). Strategies of female mate choice: a theoretical analysis. *Behavioral Ecology and Sociobiology* **7**, 107–112.

JANICKE, T. & MORROW, E. H. (2018). Operational sex ratio predicts the opportunity and direction of sexual selection across animals. *Ecology Letters* **21**, 384–391.

JANIF, Z. J., BROOKS, R. C. & DIXSON, B. J. (2014). Negative frequency-dependent preferences and variation in male facial hair. *Biology Letters* **10**, 20130958.

JENNIONS, M., KOKKO, H. & KLUG, H. (2012). The opportunity to be misled in studies of sexual selection. *Journal of Evolutionary Biology* **25**, 591–598.

JENNIONS, M. D. & PETRIE, M. (1997). Variation in mate choice and mating preferences: a review of causes and consequences. *Biological Reviews* **72**, 283–327.

JESWIET, S. B. & GODIN, J.-G. J. (2011). Validation of a method for quantifying male mating preferences in the guppy (*Poecilia reticulata*). *Ethology* **117**, 422–429.

JOHNSTONE, R. A., REYNOLDS, J. D. & DEUTSCH, J. C. (1996). Mutual mate choice and sex differences in choosiness. *Evolution***,** **50**, 1382–1391.

JORDAN, L. A. & BROOKS, R. C. (2012). Recent social history alters male courtship preferences. *Evolution* **66**, 280–287.

KACELNIK, A., VASCONCELOS, M., MONTEIRO, T. & AW, J. (2011). Darwin's "tug-of-war" vs. starlings'"horse-racing": how adaptations for sequential encounters drive simultaneous choice. *Behavioral Ecology and Sociobiology* **65**, 547–558.

KAWANO, Y. & HENMI, Y. (2016). Female preference for large waving claws in the dotillid crab *Ilyoplax pusilla*. *Journal of Ethology* **34**, 255–261.

KELLEY, L. A. & KELLEY, J. L. (2014). Animal visual illusion and confusion: the importance of a perceptual perspective. *Behavioral Ecology* **25**, 450–463.

KELLY, C. D. (2018). The causes and evolutionary consequences of variation in female mate choice in insects: the effects of individual state, genotypes and environments. *Current Opinion in Insect Science* **27**, 1–8.

KILKENNY, C., BROWNE, W. J., CUTHILL, I. C., EMERSON, M. & ALTMAN, D. G. (2010). Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *Plos Biology* **8**, e1000412.

KILMER, J., FOWLER- FINN, K., GRAY, D., HÖBEL, G., REBAR, D., REICHERT, M. & RODRÍGUEZ, R. L. (2017). Describing mate preference functions and other function-valued traits. *Journal of Evolutionary Biology* **30**, 1658–1673.

KING, S. L. (2015). You talkin'to me? Interactive playback is a powerful yet underused tool in animal communication research. *Biology Letters* **11**, 20150403.

KINGSTON, J. J., ROSENTHAL, G. G. & RYAN, M. J. (2003). The role of sexual selection in maintaining a colour polymorphism in the pygmy swordtail, *Xiphophorus pygmaeus*. *Animal Behaviour* **65**, 735–743.

KIRKPATRICK, M. (1982). Sexual selection and the evolution of female choice. *Evolution* **36**, 1–12.

KODRIC-BROWN, A. (1993). Female choice of multiple male criteria in guppies: interacting effects of dominance, coloration and courtship. *Behavioral Ecology and Sociobiology* **32**, 415–420.

KODRIC-BROWN, A. & JOHNSON, S. C. (2002). Ultraviolet reflectance patterns of male guppies enhance their attractiveness to females. *Animal Behaviour* **63**, 391–396.

KOKKO, H., BROOKS, R., JENNIONS, M. D. & MORLEY, J. (2003). The evolution of mate choice and mating biases. *Proceedings of the Royal Society B: Biological Sciences* **270**, 653–664.

KOKKO, H. & JENNIONS, M. D. (2015). Describing mate choice in a biased world. *Behavioral Ecology* **26**, 320–321.

KOKKO, H. & RANKIN, D. J. (2006). Lonely hearts or sex in the city? Density-dependent effects in mating systems. *Philosophical Transactions of the Royal Society B: Biological Sciences* **361**, 319–334.

KROODSMA, D. E., BYERS, B. E., GOODALE, E., JOHNSON, S. & LIU, W. C. (2001). Pseudoreplication in playback experiments, revisited a decade later. *Animal Behaviour* **61**, 1029–1033.

LAFLEUR, D. L., LOZANO, G. & SCLAFANI, M. (1997). Female mate-choice copying in guppies, *Poecilia reticulata*: a re-evaluation. *Animal Behaviour* **54**, 579–586.

LAHTI, D. C. (2015). The limits of artificial stimuli in behavioral research: the umwelt gamble. *Ethology* **121**, 529–537.

LANDE, R. (1981). Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences* **78**, 3721–3725.

LEINDERS-ZUFALL, T., LANE, A. P., PUCHE, A. C., MA, W., NOVOTNY, M. V., SHIPLEY, M. T. & ZUFALL, F. (2000). Ultrasensitive pheromone detection by mammalian vomeronasal neurons. *Nature* **405**, 792–796.

LENTON, A. P., FASOLO, B. & TODD, P. M. (2009). The relationship between number of potential mates and mating skew in humans. *Animal Behaviour* **77**, 55–60.

LENTON, A. P. & FRANCESCONI, M. (2011). Too much of a good thing? Variety is confusing in mate choice. *Biology Letters* **7**, rsbl20110098.

LEWIS, S. M. & CRATSLEY, C. K. (2008). Flash signal evolution, mate choice, and predation in fireflies. *Annual Reviews of Entomology* **53**, 293–321.

LUCON-XICCATO, T., SANTACÀ, M., PETRAZZINI, M. E. M., AGRILLO, C. & DADDA, M. (2019). Guppies, *Poecilia reticulata*, perceive a reversed Delboeuf illusion. *Animal Cognition* **22**, 291–303.

LUDWIG, Y. S. & POLLET, T. V. (2014). When men appear smaller or larger than they really are: preliminary evidence that women are fooled by size illusions in attractiveness judgment tasks. *Anthropological Review* **77**, 299–329.

LYNCH, K. S., RAND, A. S., RYAN, M. J. & WILCZYNSKI, W. (2005). Plasticity in female mate choice associated with changing reproductive states. *Animal Behaviour* **69**, 689–699.

LYNN, J. C. & COLE, G. L. (2019). The effect of against-background contrast on female preferences for a polymorphic colour sexual signal. *Animal Behaviour* **150**, 1–13.

MacLaren, R. D. & Rowland, W. J. (2006). Differences in female preference for male body size in *Poecilia latipinna* using simultaneous versus sequential stimulus presentation designs. *Behaviour* **143**, 273–292.

Magnhagen, C. (1991). Predation risk as a cost of reproduction. *Trends in Ecology & Evolution* **6**, 183–186.

Martel, V. & Boivin, G. (2011). Do choice tests really test choice? *Journal of Insect Behavior* **24**, 329–336.

Martin, P. & Bateson, P. (2007). *Measuring behaviour: an introductory guide*. Cambridge University Press, Cambridge.

Mattey, S. N. & Smiseth, P. T. (2015). No inbreeding avoidance by female burying beetles regardless of whether they encounter males simultaneously or sequentially. *Ethology* **121**, 1031–1038.

Mautz, B. S., Wong, B. B., Peters, R. A. & Jennions, M. D. (2013). Penis size interacts with body shape and height to influence male attractiveness. *Proceedings of the National Academy of Sciences* **110**, 6925–6930.

McGregor, P. K., Catchpole, C. K., Dabelsteen, T., Falls, J. B., Fusani, L., Gerhardt, H. C., Gilbert, F., Horn, A. G., Klump, G. M. & Kroodsma, D. E. (1992). Design of playback experiments: the Thornbridge Hall NATO ARW consensus. In *Playback and studies of animal communication* (ed. P. K. McGregor), pp. 1–9. Springer, Boston.

McNamara, K. B., Jones, T. M. & Elgar, M. A. (2004). Female reproductive status and mate choice in the hide beetle, *Dermestes maculatus*. *Journal of Insect Behavior* **17**, 337–352.

Meuche, I., Brusa, O., Linsenmair, K. E., Keller, A. & Pröhl, H. (2013). Only distance matters–non-choosy females in a poison frog population. *Frontiers in Zoology* **10**, 29.

MILINSKI, M. & BAKKER, T. C. (1990). Female sticklebacks use male coloration in mate choice and hence avoid parasitized males. *Nature* **344**, 330–333.

MILINSKI, M. & BAKKER, T. C. (1992). Costs influence sequential mate choice in sticklebacks, *Gasterosteus aculeatus*. *Proceedings of the Royal Society B: Biological Sciences* **250**, 229–233.

MILLER, M. N. & FINCKE, O. M. (1999). Cues for mate recognition and the effect of prior experience on mate recognition in Enallagma damselflies. *Journal of Insect Behavior* **12**, 801–814.

MITOYEN, C., QUIGLEY, C. & FUSANI, L. (2019). Evolution and function of multimodal courtship displays. *Ethology* **125**, 503–515.

MORAVEC, M. L., STRIEDTER, G. F. & BURLEY, N. T. (2010). ‘Virtual parrots’ confirm mating preferences of female budgerigars. *Ethology* **116**, 961–971.

MORGAN, K. V., HURLY, T. A., MARTIN, L. & HEALY, S. D. (2016). Presentation order affects decisions made by foraging hummingbirds. *Behavioral Ecology and Sociobiology* **70**, 21–26.

MURAYAMA, T., USUI, A., TAKEDA, E., KATO, K. & MAEJIMA, K. (2012). Relative size discrimination and perception of the Ebbinghaus Illusion in a bottlenose dolphin (*Tursiops truncatus*). *Aquatic Mammals* **38**, 333–342.

NIEBERDING, C. M. & HOLVECK, M.-J. (2017). Laboratory social environment biases mating outcome: a first quantitative synthesis in a butterfly. *Behavioral Ecology and Sociobiology* **71**, 117.

NITYANANDA, V. (2016). Attention-like processes in insects. *Proceedings of the Royal Society B: Biological Sciences* **283**, 20161986.

OLIVEIRA, R. F., ROSENTHAL, G. G., SCHLUPP, I., MCGREGOR, P. K., CUTHILL, I. C., ENDLER, J. A., FLEISHMAN, L. J., ZEIL, J., BARATA, E. & BURFORD, F. (2000). Considerations on the use of video playbacks as visual stimuli: the Lisbon workshop consensus. *Acta Ethologica* **3**, 61–65.

OWEN, M. A., ROHRER, K. & HOWARD, R. D. (2012). Mate choice for a novel male phenotype in zebrafish, *Danio rerio*. *Animal Behaviour* **83**, 811–820.

PARRI, S., ALATALO, R. V., KOTIAHO, J. & MAPPES, J. (1997). Female choice for male drumming in the wolf spider *Hygrolycosa rubrofasciata*. *Animal Behaviour* **53**, 305–312.

PARRON, C. & FAGOT, J. (2007). Comparison of grouping abilities in humans (*Homo sapiens*) and baboons (*Papio papio*) with the Ebbinghaus illusion. *Journal of Comparative Psychology* **121**, 405–411.

PETRIE, M. & HALLIDAY, T. (1994). Experimental and natural changes in the peacock's (*Pavo cristatus*) train can affect mating success. *Behavioral Ecology and Sociobiology* **35**, 213–217.

PFENNIG, K. S. (2007). Facultative mate choice drives adaptive hybridization. *Science* **318**, 965–967.

PHELPS, S. M., RAND, A. S. & RYAN, M. J. (2006). A cognitive framework for mate choice and species recognition. *The American Naturalist* **167**, 28–42.

PILAKOUTA, N. & SMISETH, P. T. (2017). Female mating preferences for outbred versus inbred males are conditional upon the female's own inbreeding status. *Animal Behaviour* **123**, 369–374.

PITCHER, T. E., NEFF, B. D., RODD, F. H. & ROWE, L. (2003). Multiple mating and sequential mate choice in guppies: females trade up. *Proceedings of the Royal Society B: Biological Sciences* **270**, 1623–1629.

PIZZARI, T. & BIRKHEAD, T. (2000). Female feral fowl eject sperm of subdominant males. *Nature* **405**, 787–789.

PLATH, M., SEGGEL, U., BURMEISTER, H., HEUBEL, K. U. & SCHLUPP, I. (2006). Choosy males from the underground: male mating preferences in surface-and cave-dwelling Atlantic mollies (*Poecilia mexicana*). *Naturwissenschaften* **93**, 103–109.

POLLACK, G. S. (1988). Selective attention in an insect auditory neuron. *Journal of Neuroscience* **8**, 2635–2639.

POWELL, D. L. & ROSENTHAL, G. G. (2016). What artifice can and cannot tell us about animal behavior. *Current Zoology* **63**, 21–26.

RAFFA, K. F., HAVILL, N. P. & NORDHEIM, E. V. (2002). How many choices can your test animal compare effectively? Evaluating a critical assumption of behavioral preference tests. *Oecologia* **133**, 422–429.

REANEY, L. T., SIMS, R. A., SIMS, S. W., JENNIONS, M. D. & BACKWELL, P. R. (2008). Experiments with robots explain synchronized courtship in fiddler crabs. *Current Biology* **18**, R62–R63.

REICHERT, M. S. & HÖBEL, G. (2015). Modality interactions alter the shape of acoustic mate preference functions in gray treefrogs. *Evolution* **69**, 2384–2398.

REINHOLD, K. & SCHIELZETH, H. (2015). Choosiness, a neglected aspect of preference functions: a review of methods, challenges and statistical approaches. *Journal of Comparative Physiology A* **201**, 171–182.

RITCHIE, M. G. (1996). The shape of female mating preferences. *Proceedings of the National Academy of Sciences* **93**, 14628–14631.

RITCHIE, M. G. (2007). Sexual selection and speciation. *Annual Review of Ecology, Evolution and Systematics* **38**, 79–102.

RODRÍGUEZ, R. L., HALLETT, A., KILMER, J. & FOWLER- FINN, K. (2013*a*). Curves as traits: genetic and environmental variation in mate preference functions. *Journal of Evolutionary Biology* **26**, 434–442.

RODRÍGUEZ, R. L., REBAR, D. & FOWLER-FINN, K. D. (2013*b*). The evolution and evolutionary consequences of social plasticity in mate preferences. *Animal Behaviour* **85**, 1041–1047.

ROFF, D. & FAIRBAIRN, D. (2015). Bias in the heritability of preference and its potential impact on the evolution of mate choice. *Heredity* **114**, 404–412.

ROFF, D. A. & FAIRBAIRN, D. J. (2014). The evolution of phenotypes and genetic parameters under preferential mating. *Ecology and Evolution* **4**, 2759–2776.

RÖMER, H. & KRUSCH, M. (2000). A gain-control mechanism for processing of chorus sounds in the afferent auditory pathway of the bushcricket *Tettigonia viridissima* (Orthoptera; Tettigoniidae). *Journal of Comparative Physiology A* **186**, 181–191.

RONALD, K. L., FERNANDEZ-JURICIC, E. & LUCAS, J. R. (2012). Taking the sensory approach: how individual differences in sensory perception can influence mate choice. *Animal Behaviour* **84**, 1283–1294.

RONALD, K. L., FERNÁNDEZ-JURICIC, E. & LUCAS, J. R. (2018). Mate choice in the eye and ear of the beholder? Female multimodal sensory configuration influences her preferences. *Proceedings of the Royal Society B: Biological Sciences* **285**, 20180713.

ROSENTHAL, G. G. (2017). *Mate choice*. Princeton University Press, Princeton.

ROSENTHAL, G. G. & RYAN, M. J. (2000). Visual and acoustic communication in non-human animals: a comparison. *Journal of Biosciences* **25**, 285–290.

ROWLAND, W. J. (1982). Mate choice by male sticklebacks, *Gasterosteus aculeatus*. *Animal Behaviour* **30**, 1093–1098.

ROYLE, N. J., LINDSTRÖM, J. & METCALFE, N. B. (2008). Context-dependent mate choice in relation to social composition in green swordtails *Xiphophorus helleri*. *Behavioral Ecology* **19**, 998–1005.

RYAN, M., AKRE, K. & KIRKPATRICK, M. (2009). Cognitive mate choice. In *Cognitive ecology II* (ed. R. Dukas and J. M. Ratcliffe), pp. 137–155. The University of Chicago Press, Chicago.

RYAN, M. J. & CUMMINGS, M. E. (2013). Perceptual biases and mate choice. *Annual Review of Ecology, Evolution, and Systematics* **44**, 437–459.

RYAN, M. J. & KEDDY-HECTOR, A. (1992). Directional patterns of female mate choice and the role of sensory biases. *The American Naturalist* **139**, S4–S35.

RYAN, M. J. & TAYLOR, R. C. (2015). Measures of mate choice: comment on Dougherty and Shuker. *Behavioral Ecology* **26**, 323–324.

SANDVIK, M., ROSENQVIST, G. & BERGLUND, A. (2000). Male and female mate choice affects offspring quality in a sex–role–reversed pipefish. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **267**, 2151–2155.

SCHEIBEHENNE, B., GREIFENEDER, R. & TODD, P. M. (2010). Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research* **37**, 409–425.

SCHWARTZ, B. (2004). *The paradox of choice: Why more is less*. Ecco, New York.

SCHWARTZ, J. J., BUCHANAN, B. W. & GERHARDT, H. C. (2001). Female mate choice in the gray treefrog (*Hyla versicolor*) in three experimental environments. *Behavioral Ecology and Sociobiology* **49**, 443–455.

SCHWARTZ, J. J., HUTH, K. & HUTCHIN, T. (2004). How long do females really listen? Assessment time for female mate choice in the grey treefrog, *Hyla versicolor*. *Animal Behaviour* **68**, 533–540.

SEEHAUSEN, O., VAN ALPHEN, J. J. & WITTE, F. (1997). Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* **277**, 1808-1811.

SENAR, J., MATEOS-GONZALEZ, F., URIBE, F. & ARROYO, L. (2013). Familiarity adds to attractiveness in matters of siskin mate choice. *Proceedings of the Royal Society B: Biological Sciences* **280**, 20132361.

SHUKER, D. M. & DAY, T. H. (2001). The repeatability of a sexual conflict over mating. *Animal Behaviour* **61**, 755–762.

SIMMONS, L. & BAILEY, W. (1990). Resource influenced sex roles of zaprochiline tettigoniids (Orthoptera: Tettigoniidae). *Evolution* **44**, 1853–1868.

SIMMONS, L. W. (2001). *Sperm competition and its evolutionary consequences in the insects*. Princeton University Press, Princeton.

SLAGSVOLD, T., HANSEN, B. T., JOHANNESSEN, L. E. & LIFJELD, J. T. (2002). Mate choice and imprinting in birds studied by cross-fostering in the wild. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **269**, 1449–1455.

SPIETH, H. T. (1974). Courtship behavior in Drosophila. *Annual Review of Entomology* **19**, 385–405.

ST JOHN, M. E. & FULLER, R. C. (2018). The effects of experimental design on mating preferences and reproductive isolation in killifish. *Behavioral Ecology* **30**, 92–100.

STODDARD, M. C., MILLER, A. E., EYSTER, H. N. & AKKAYNAK, D. (2018). I see your false colours: how artificial stimuli appear to different animal viewers. *Journal of The Royal Society Interface* **9**, 20180053.

SULLIVAN, M. S. (1994). Mate choice as an information gathering process under time constraint: implications for behaviour and signal design. *Animal Behaviour* **47**, 141–151.

SVENSSON, P. A., LEHTONEN, T. K. & WONG, B. B. M. (2010). The interval between sexual encounters affects male courtship tactics in a desert-dwelling fish. *Behavioral Ecology and Sociobiology* **64**, 1967–1970.

SWADDLE, J. P. & PAGE, L. C. (2007). High levels of environmental noise erode pair preferences in zebra finches: implications for noise pollution. *Animal Behaviour* **74**, 363–368.

TALLAMY, D. W., POWELL, B. E. & MCCLAFFERTY, J. A. (2002). Male traits under cryptic female choice in the spotted cucumber beetle (Coleoptera: Chrysomelidae). *Behavioral Ecology* **13**, 511–518.

TEN CATE, C. & VOS, D. R. (1999). Sexual imprinting and evolutionary processes in birds: a reassessment. In *Advances in the Study of Behavior*, vol. 28 (ed. P. B. Slater, J. S. Rosenblat, C. T. Snowden and T. J. Roper), pp. 1–31. Academic Press, San Diego.

UETZ, G. W., ROBERTS, J. A. & TAYLOR, P. W. (2009). Multimodal communication and mate choice in wolf spiders: female response to multimodal versus unimodal signals. *Animal Behaviour* **78**, 299–305.

VAKIRTZIS, A. (2011). Mate choice copying and nonindependent mate choice: a critical review. *Annales Zoologici Fennici* **48**, 91–108.

VALONE, T. J., NORDELL, S. E., GIRALDEAU, L.-A. & TEMPLETON, J. J. (1996). The empirical question of thresholds and mechanisms of mate choice. *Evolutionary Ecology* **10**, 447–455.

VAN GOSSUM, H., STOKS, R. & BRUYN, L. D. (2001). Reversible frequency–dependent switches in male mate choice. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **268**, 83–85.

VERSPOOR, R., CUSS, M. & PRICE, T. (2015). Age-based mate choice in the monandrous fruit fly *Drosophila subobscura*. *Animal Behaviour* **102**, 199–207.

VITOUSEK, M. N., MITCHELL, M. A., WOAKES, A. J., NIEMACK, M. D. & WIKELSKI, M. (2007). High costs of female choice in a lekking lizard. *Plos One* **2**, e567.

VUKOMANOVIC, J. & RODD, F. H. (2007). Size- dependent female mate copying in the guppy (*Poecilia reticulata*): large females are role models but small ones are not. *Ethology* **113**, 579–586.

WAGNER, JR W. E. (1998). Measuring female mating preferences. *Animal Behaviour* **55**, 1029–1042.

WAGNER, JR W. E. & HARPER, C. J. (2003). Female life span and fertility are increased by the ejaculates of preferred males. *Evolution* **57**, 2054–2066.

WAGNER, JR W. E., MURRAY, A.-M. & CADE, W. H. (1995). Phenotypic variation in the mating preferences of female field crickets, *Gryllus integer*. *Animal Behaviour* **49**, 1269–1281.

WALLING, C. A., ROYLE, N. J., LINDSTROM, J. & METCALFE, N. B. (2010). Do female association preferences predict the likelihood of reproduction? *Behavioral Ecology and Sociobiology* **64**, 541–548.

WERNER, N. Y. & LOTEM, A. (2006). Experimental evidence for male sequential mate preference in a lekking species. *Ethology* **112**, 657–663.

WHITE, D. J. & GALEF JR, B. G. (1999). Affiliative preferences are stable and predict mate choices in both sexes of Japanese quail, *Coturnix japonica*. *Animal Behaviour* **58**, 865–871.

WIDEMO, F. & SAETHER, S. A. (1999). Beauty is in the eye of the beholder: causes and consequences of variation in mating preferences. *Trends in Ecology & Evolution* **14**, 26–31.

WILEY, R. H. & POSTON, J. (1996). Perspective: indirect mate choice, competition for mates, and coevolution of the sexes. *Evolution* **50,** 1371–1381.

WILLIS, P. M., RYAN, M. J. & ROSENTHAL, G. G. (2011). Encounter rates with conspecific males influence female mate choice in a naturally hybridizing fish. *Behavioral Ecology* **22**, 1234–1240.

WITHERS, T. & MANSFIELD, S. (2005). Choice or no-choice tests? Effects of experimental design on the expression of host range. In *Second International Symposium on Biological Control of Arthropods,* pp. 620–633. United States Department of Agriculture, Forest Service, Davos, Switzerland.

WITTE, K. (2006). Time spent with a male is a good indicator of mate preference in female zebra finches. *Ethology Ecology & Evolution* **18**, 195–204.

WONG, B. & CANDOLIN, U. (2005). How is female mate choice affected by male competition? *Biological Reviews* **80**, 559–571.

YORZINSKI, J. L., PATRICELLI, G. L., BABCOCK, J. S., PEARSON, J. M. & PLATT, M. L. (2013). Through their eyes: selective attention in peahens during courtship. *Journal of Experimental Biology* **216**, 3035–3046.

ZAJITSCHEK, S. R. & BROOKS, R. C. (2008). Distinguishing the effects of familiarity, relatedness, and color pattern rarity on attractiveness and measuring their effects on sexual selection in guppies (*Poecilia reticulata*). *The American Naturalist* **172**, 843–854.

ZIEGE, M., HENNIGE-SCHULZ, C., MUECKSCH, F., BIERBACH, D., TIEDEMANN, R., STREIT, B. & PLATH, M. (2012). A comparison of two methods to assess audience-induced changes in male mate choice. *Current Zoology* **58**, 84–94.

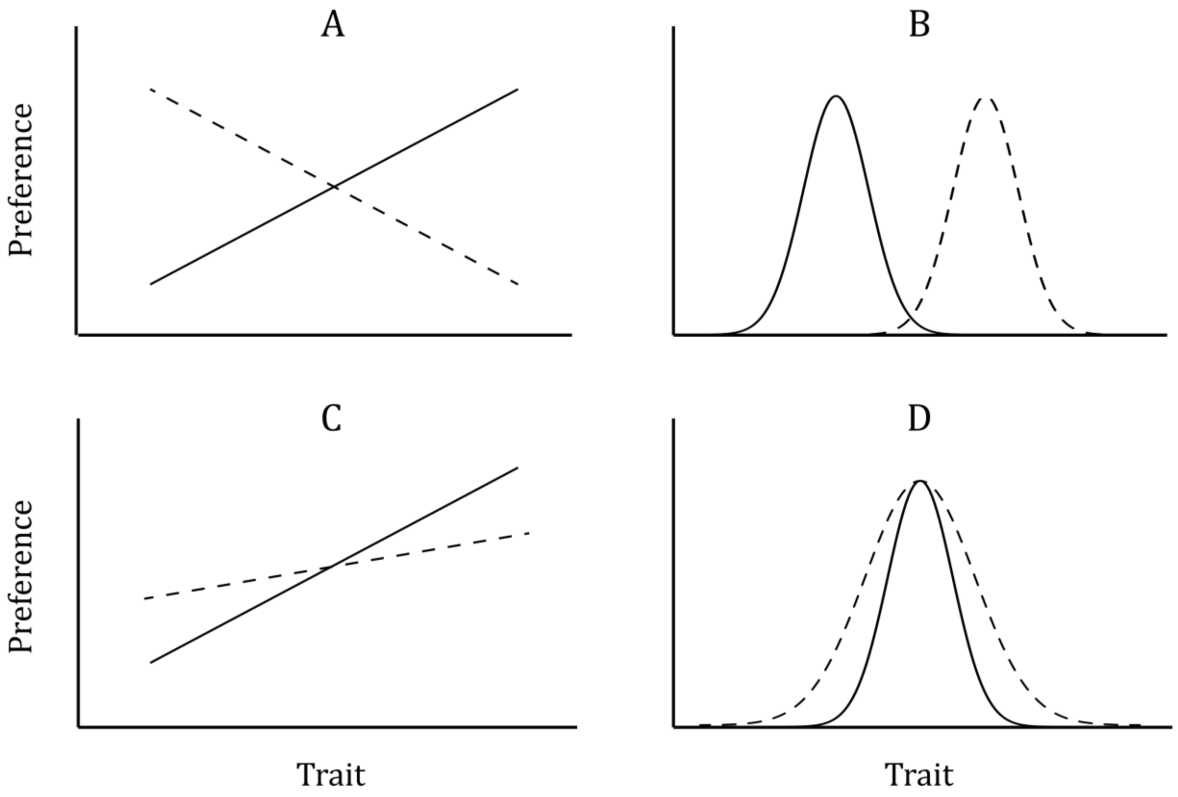

Fig. 1. Mating preferences can be described using a preference function, which is obtained by plotting the relationship between a mate phenotype and chooser preference. Two key ways that preference functions can vary are in relation to the peak preference (A, B) and the strength of preference (C, D). Change in the peak preference can lead to a change in the direction of a linear preference (from a positive to a negative slope; A), or a horizontal shift in a stabilising preference (B). Increasing the strength of preference (C, D; change from the broken line to the solid line) leads to a steeper slope in the case of a linear preference C), or a tighter curve in the case of a quadratic preference (D).

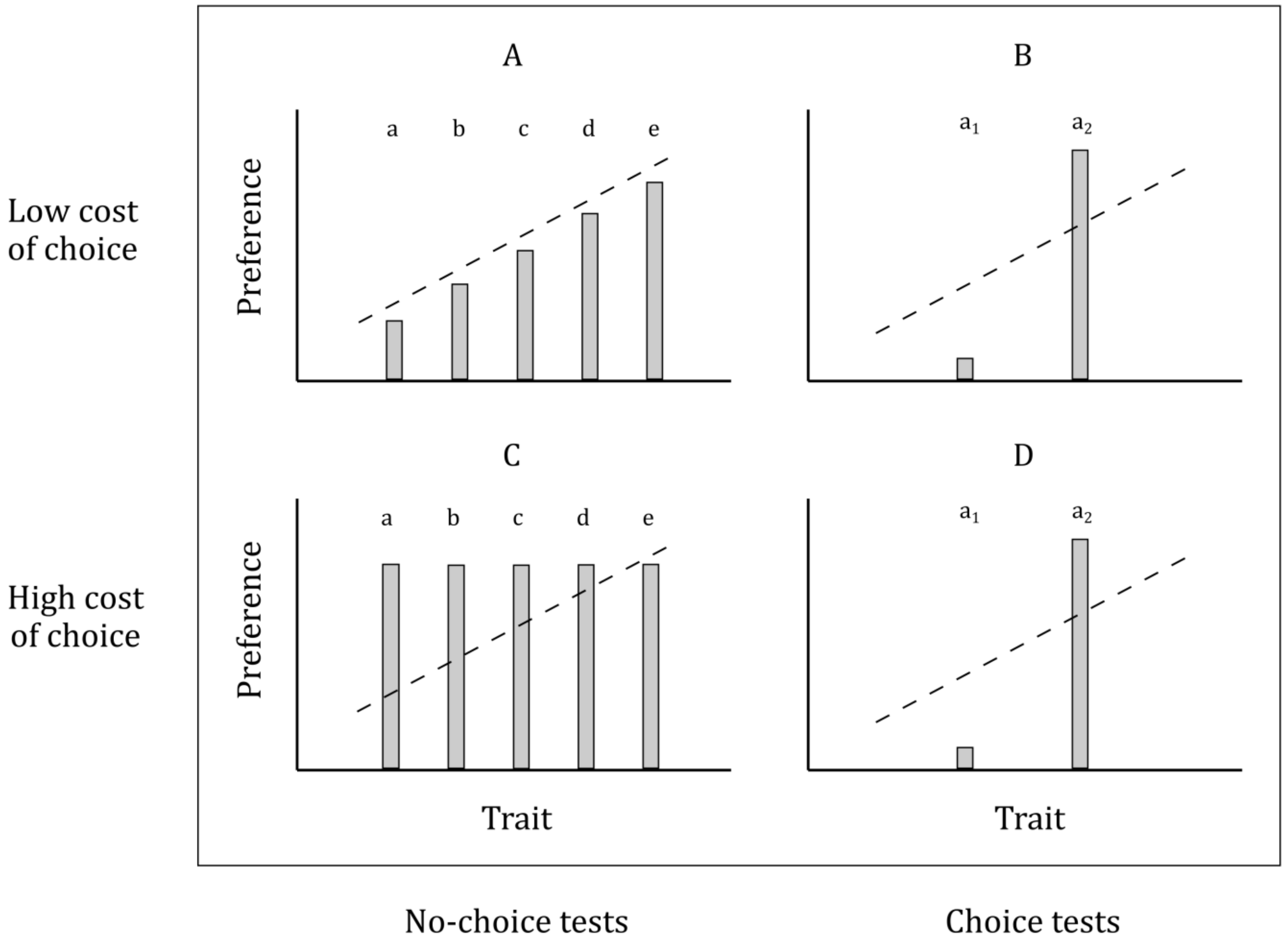

Fig. 2. Two examples showing how the different choice designs may under- or overestimate preferences. In all cases, the dotted line represents the underlying preference function we are trying to measure, and the bars represent the actual recorded choices (either at the individual or population level). The letters above the bars illustrate the different trials used to test for preferences: for no-choice tests, each letter represents a

different trial; for choice tests, a_1 and a_2 represent the two choice options presented during a single, two-choice trial. First, when the costs of choice are low (A, B), choice tests may exaggerate the strength of preference when a dichotomous (all-or-nothing) measure of preference is used, because they test relative preferences (B). This means that no matter how small the difference in preference between options, the observed preference will be strong. For a dichotomous preference, no-choice outcomes should match the underlying preference well (A). Second, when the costs of choice are high (C, D), subjects in no-choice tests may respond maximally to all options, or mate randomly with respect to any given trait, because the perceived risk of remaining unmated is high. This will lead to no preference being detected (C). Therefore, for no-choice tests the strength of preference will be influenced by the costs of choice. Preferences in choice tests should not be influenced by the costs of choice (D), because rejection of one option is typically cost-free.

Table 1. Outline of strengths and potential problems associated with the main experimental approaches used in animal mate choice experiments, in relation to the five design considerations covered in this review.

Consideration	***Section***	***Category***	***Strengths***	***Potential problems***
How is mate choice measured?	III	Mating outcomes	Directly reflects fitness	Mating may influence receptivity Difficulty determining causal link between trait and preference Difficulty attributing choice to either sex Binary outcome reduces statistical power
		Proxy measure	More experimental designs possible Can determine causal link between trait and preference Can attribute choice to one sex Allows for continuous measures of preference	Scoring of behaviour can be subjective Less obvious which behaviour to measure Behaviours may not reflect choice outcomes
Are subjects given a choice during tests?	IV	No-choice	No intrasexual interactions between stimuli Test absolute preferences	May exaggerate cost of rejecting options Comparative evaluation by choosers not possible
		Choice	Can test multiple stimuli at once Comparative evaluation by choosers is possible Test relative preferences	Intrasexual interactions between stimuli may confound choice Non-responsive subjects are often excluded from analyses Two-choice tests can exaggerate preferences
How many times is each subject tested?	V	Single test	No effect of past experience Can test more individuals	Cannot detect complex individual preferences
		Multiple tests	Can detect complex individual preferences Can detect repeatability of preference	Experience effects in later trials Interval between presentations alters mate encounter rate
How many options are available during choice tests?	VI	Two	See choice category above	See choice category above
		More than two	Fewer trials needed with each subject Ecological realism	Subjects may not perceive or assess all stimuli Stimuli may interfere with each other Choosers may take longer to choose

What stimuli are subjects required to choose between?	VII	Live animals	Realism	Difficulty determining causal link between trait and preference Courter responses may influence chooser behaviour More animals needed for experiments
		Natural stimuli	Fewer animals used during experiments Can determine causal link between trait and preference	May not be perceived as signals Prone to pseudoreplication
		Artificial stimuli	Fewer animals used during experiments Can determine causal link between trait and preference Can test preference for novel traits	May not be perceived as signals Prone to pseudoreplication

Table 2. Example studies that have tested whether proxy behavioural measures of preference correlate with choice outcomes (or another mating behaviour in species for which mating is difficult to observe).

Group	***Species***	***Study***	***Proxy measure***	***Choice measure***	***Correlated measures?***
Fish	*Danio rerio*	Owen *et al*. (2012)	Association time	Mating latency	No
	Poecilia reticulata	Kodric-Brown (1993)	Association time	Mating	Yes
		Jeswiet & Godin (2011)	Association time	Mating attempts	No
	Poecilia mexicana	Plath *et al*. (2006)	Association time	Mating attempts[1]	Yes
		Ziege *et al*. (2012)	Association time	Mating attempts[1]	Yes
	Pseudotropheus zebra	Couldridge & Alexander (2001)	Association time	Mating attempts[2]	Yes
	Salaria pavo	Goncalves & Oliveira (2003)	Association time	Mating	No
			Courtship displays	Mating	No
	Xiphophorus helleri	Walling *et al*. (2010)	Association time	Mating	Yes
Birds	*Carduelis spinus*	Senar *et al*. (2013)	Association time	Courtship feeding	Yes
	Coturnix japonica	White & Galef (1999)	Association time	Mating	Yes
Mammals	*Peromyscus californicus*	Gubernick & Addington (1994)	Association time[3]	Mating	No

[1] Includes both successful and unsuccessful mating attempts (sperm transfer not measured).

[2] Mating attempts when sexes separated by glass partition.

[3] Association time measured when females were not in oestrus.

Table 3. Example studies which have tested mating preferences using both no-choice and choice tests. In the majority of cases, preferences are stronger in choice tests compared to no-choice tests. This difference cannot be attributed to intrasexual competition between choice options in choice tests: in most cases rivals cannot interact directly.

Group	***Species***	***Study***	***Chooser sex***	***Chosen trait***	***Preference observed in***	***Preference stronger in choice tests?***	***Could options interact?***
Arachnids	*Hygrolycosa rubrofasciata*	Parri *et al*. (1997)	Female	Drumming volume	Both	No	No
Crustaceans	*Uca mjoebergi*	Booksmythe *et al*. (2011)	Male	Species	Choice only	–	No
Insects	*Cadra cautella*	Allison & Cardé (2008)	Male	Pheromone blend	Choice only	–	No
	Dermestes maculatus	McNamara *et al*. (2004)	Male	Mated status	Choice only	–	No
	Drosophila melanogaster	Baxter *et al*. (2018)	Female	Body size	Both	No	Yes
			Female	Age	Both	Yes	Yes
	Drosophila subobscura	Verspoor *et al*. (2015)	Female	Age	Both	No	Yes
	Drosophila santomea	Coyne *et al*. (2005)	Female	Species	Both	No	Yes
			Male	Species	Both	Yes	Yes
	Drosophila yakuba		Female	Species	Both	Yes	Yes
			Male	Species	Both	No	Yes
	Gryllus integer	Wagner *et al*. (1995)	Female	Song	Both	Yes	No
	Nicrophorus vespilloides	Mattey & Smiseth (2015)	Female	Relatedness	Neither	–	No
	Pseudomantis albofimbriata	Barry *et al*. (2010)	Male	Condition	Choice only	–	No
Fish	*Danio rerio*	Owen *et al*. (2012)	Female	Colour	Choice only	–	No
	Gambusia holbrooki	Head *et al*. (2015)	Male	Body size	Choice only	–	Yes
	Gasterosteus aculeatus	Rowland (1982)	Male	Body size	Choice only	–	No
	Lucania goodei	St John & Fuller (2018)	Male	Species	Both	Yes	No

			Female	Species	No-choice only	–	No
	Poecilia latipinna	MacLaren & Rowland (2006)	Female	Body size	Both	Yes	No
	Poecilia reticulata	Jordan & Brooks (2012)	Male	Body size	Both	No	No
	Xiphophorus birchmanni	Willis *et al*. (2011)	Female	Species	Choice only	–	No
Amphibians	*Notophthalmus viridescens*	Gabor *et al*. (2000)	Female	Tail height	Neither	–	No
	Physalaemus pustulosus	Phelps *et al*. (2006)	Female	Song	Both	Yes	No
Birds	*Gallus gallus*	Gillingham *et al*. (2008)	Female	Genetic dissimilarity	Neither	-	No

Table 4. Recommendations for the best approach for the accurate reconstruction of a preference function. The number of stimuli needed, and the number of times each individual needs to be tested, depend on the shape of the preference function being measured, the level of preference considered (individual or population level), and the choice design being used.

Preference shape	**Preference level**	**Number of stimuli needed**	**Times individuals tested** **No-choice test**	 **Choice test**
Linear	Individual	*Two*	*At least twice*	*Once*
	Population	*Two*	*Once*	*Once*
Complex	Individual	*More than two*	*More than twice*	*Once or more*
	Population	*More than two*	*Once*	*Once or more*