Behavioral Ecology

The official journal of the
ISBE
International Society for Behavioral Ecology

Behavioral Ecology (2018), XX(XX), 1–10. doi:10.1093/beheco/ary150

Original Article

The effects of experimental design on mating preferences and reproductive isolation in killifish

Michelle E. St. John and Rebecca C. Fuller
School of Integrative Biology, University of Illinois Urbana – Champaign, Champaign, IL 61820, USA

Received 3 May 2018; revised 26 September 2018; editorial decision 2 October 2018; accepted 16 October 2018.

Determining the direction and magnitude of mating preference is fundamental to many questions in evolutionary biology. Unlike the measurement of traits such as body size, the measurement of mating preferences is likely affected by experimental design. Scientists must choose both the behavioral assay in which to measure preference and the metrics that serve as a proxy for preference. The accuracy of these assays and metrics, however, is often unknown and seldom tested. Here, we compared the accuracy of 3 assays (dichotomous choice, audience assay, and no-choice assay) and 3 metrics (association time, courtship bouts, and number of eggs produced) in the bluefin killifish, which possesses strong, conspecific mating preferences when in sympatry with rainwater killifish. We consistently detected preferences in both males and females when using metrics associated with mating (i.e., courting bouts and number of eggs spawned). However, we failed to consistently detect preference when using time as a metric. We then used all 3 assays and metrics to test for cascade reinforcement. Cascade reinforcement predicts that enhanced behavioral isolation between sympatric species creates enhanced behavioral isolation among populations within species. We tested whether male and female bluefin killifish had heightened preference for mates from native over foreign populations. We consistently detected female preferences for native males, but did not detect male preferences for native females. Reproductive isolation values also reflect these preferences. Ultimately, we illustrated the importance of using multiple approaches to evaluate and legitimize measures of mating preference for males and females choosing among mates in different contexts.

Key words: **cascade reinforcement, dichotomous choice, female mate choice, male mate choice, no-choice assay, reinforcement.**

INTRODUCTION

Many questions in evolutionary biology require the measurement of animal mate preference. Accurately determining mate preference in the laboratory, however, is quite challenging. There are 2 critical considerations for laboratory assays of mate preference. The first is how a focal individual is presented with stimulus mates (hereafter referred to as the "assay"). The second is the behaviors or actions that are measured during the assay as a proxy for mate preference (hereafter referred to as the "metric"). Generally, behavioral assays can be broken into 2 categories: no-choice assays and choice assays (Dougherty and Shuker 2015). While choice assays present a focal individual with 2 or more potential mates, no-choice assays present a focal individual with only a single potential mate (Rundle and Schluter 1998; Wagner 1998; McGhee et al. 2007; Nosil 2007; Dougherty and Shuker 2015; reviewed in Rosenthal (2017)). Metrics also vary and can range from condition-dependent behaviors, such as the frequency of courting bouts, to condition-independent behaviors, such as association time (Hunt et al. 2005; Cotton et al. 2006; Cummings and Mollaghan 2006). An assay's or metric's ability to accurately detect mate preference, however, can depend on many factors.

Many studies assume that their behavioral assays reliably detect the correct direction and magnitude of mate preference. An assay's or metric's ability to do this, however, can depend on whether organisms naturally encounter mates concurrently or sequentially (Dougherty and Shuker 2015; Ryan and Taylor 2015), whether mate preference is condition-dependent (Hunt et al. 2005; Cotton et al. 2006; Kokko and Jennions 2015), or even the strength of mate preference (Houde 1997; Coyne and Orr 2004). In their 2015 meta-analysis, Dougherty and Shuker found that measuring mate preference for the same species using multiple assays often resulted in differing reports of mate preference, with choice assays consistently detecting stronger mate preferences than no-choice assays. Furthermore, a single assay could vary in its ability to detect preference depending on the sex of the organism or whether preference was at the between- or within-species level (Dougherty and Shuker 2015). Metrics can similarly vary in their ability to reliably detect

Address correspondence to M.E. St. John. E-mail: michelle.e.stjohn@gmail.com

mate preference. For example, Cummings and Mollaghan (2006) measured female northern swordtail (*Xiphophorus nigrensis*) mate preference using 2 metrics: glides (a mating behavior) and association time. They found, however, that only association time was repeatable across several days. Clearly, methodology for measuring mate preference is highly variable, but few studies verify that their assays or metrics genuinely predict the outcome of mating.

Despite the fact that most studies do not investigate the reliability of the assays or metrics used to detect mate preference, a few studies do give general advice for detecting mate preference. First, Wagner (1998) suggests that measuring an individual's mate preference several times provides a more accurate description of their true preference. Second, Dougherty and Shuker (2015) suggest using several different assays to measure mate preference, as similar results will corroborate any findings. Like measures of heritability, repeatability, or selection differentials, the manifestation of mating preferences undoubtedly depends on methods used to assess them, the populations of animals investigated, and the time of year/breeding season over which preferences are measured (Charmantier et al. 2014). Robust inferences may only be possible when multiple studies using multiple methods and metrics are employed. Measuring preferences repeatedly also allows us to gauge the reliability of assays and metrics for a given system.

Here, we sought to determine which assays and metrics most reliably detect mate preference across 3 mate choice contexts in the bluefin killifish (*Lucania goodei*). *Lucania goodei* is an excellent species to test the reliability of behavioral assays and metrics for 2 reasons. First, the *L. goodei* mating ritual involves several steps and behaviors that may serve as different metrics or measures of mate preference. For example, in the wild, male *L. goodei* defend small territories which females visit—making association time a plausible metric for measuring mate preference. Similarly, males court females whom they are presumably interested in—indicating that courting may also be a good metric for mate preference. At this point, females may remain close to the male and continue to be courted or females may leave the territory. Finally, wild encounters often culminate in spawning—indicating that egg production could also be used as a metric for preference. Second, *L. goodei* has strong and well-documented conspecific mate preference when found in sympatry with their sister species, *Lucania parva* (Fuller et al. 2007; Fuller 2008; Berdan and Fuller 2012; Gregorio et al. 2012; Kozak et al. 2015). In sympatry, *L. goodei* and *L. parva* mate and produce unfit hybrids at low levels (Walker and Johnson 1943; Fuller 2008). Selection against hybridization favors increased conspecific mate preference (prezygotic isolation/behavioral isolation) in a phenomenon termed reinforcement. Reinforcement is potentially a very potent evolutionary force, as it is the only form of selection that can directly increase reproductive isolation (RI) between groups. Although initially met with skepticism, reinforcement is now widely accepted and has been documented in the *Lucania* system as well as several other species (birds: Sætre et al. (1997); frogs: Blair (1974) and Lemmon (2009); fish: Fuller et al. (2007); plants: Hopkins et al. (2014); and insects: Kelly and Noor (1996), Yukilevich and True (2006), and Nosil (2007)). The effects of reinforcement in the *Lucania* system not only make *L. goodei* an ideal study organism for testing the reliability of metrics and assays, but also allows us to test 2 additional implications of reinforcement.

First, reinforcement predicts that the increase in prezygotic isolation should coincide with increased costs of hybridization (Yukilevich 2012). This means that asymmetric hybrid fitness, between species or even between sexes, should lead to asymmetric prezygotic isolation (Hoffmann and Turelli 1997; Pfennig and Simovich 2002; Jaenike et al. 2006; Bolnick et al. 2008). As an example, consider 2 species A and B where hybrids formed by matings between A females and B males have lower fitness than the reciprocal hybrid (B females × A males). Reinforcement may result in a scenario where A females and B males are less likely to hybridize compared with B females and A males. In fact, Yukilevich (2012) found good evidence for such a pattern in *Drosophila*. Another possibility, however, is that females should always have higher levels of behavioral isolation than males. In general, females invest more than males in a given reproductive event, particularly in systems where males do not provide parental care (Wirtz 1999; Coyne and Orr 2004; Yukilevich 2012).

Lucania is a good system to test these scenarios because there are asymmetric fitness costs to hybridization for male and female *L. goodei* and *L. parva*. Crosses between *L. goodei* females and *L. parva* males produce F1-hybrid offspring with no discernible decrease in fitness (Fuller 2008). On the other hand, crosses between *L. goodei* males and *L. parva* females produce F1 males whose fertilization success is reduced by at least 50% (Fuller 2008). The hypothesis that asymmetries in hybrid fitness should be reflected in the strength of behavioral isolation predicts that *L. goodei* males should have high levels of preference for conspecifics. In contrast, the hypothesis that females energetically invest more into a given reproductive event (and that a given mating event is cheap for males) predicts that females should have high levels of conspecific preference.

The second implication is that reinforcement can lead to correlated effects on within-species preferences. Reinforcement's signature is shifted-mating traits and preferences in areas of sympatry compared with allopatry. If traits and preferences shift drastically, then individuals from sympatric populations may begin to discriminate against conspecific mates from foreign populations. This increase in native mate preference as an incidental effect of reinforcement is known as cascade reinforcement (Ortiz-Barrientos et al. 2009; Fuller 2016; Pfennig 2016) and has been documented in only 2 fish systems (*L. parva*: Kozak et al. (2015) and darters: Moran et al. (2017) and Moran and Fuller (2018)). Cascade reinforcement can theoretically result in rapid diversification of mating traits and preferences leading to speciation events in the absence of postzygotic isolation (a concept previously attributed only to sexual selection; Hoskin and Higgie 2010; Pfennig and Rice 2014), but its frequency in nature is still largely unknown. Cascade reinforcement, however, has been documented in *L. goodei*'s sister species, *L. parva*. Kozak et al. (2015) found that sympatric female *L. parva* preferred native mates significantly more than foreign mates, whereas allopatric female *L. parva* showed no preference. Whether cascade reinforcement is also present in *L. goodei* is unknown.

In summary, our experiment had 3 goals. The first goal was to determine which behavioral assays or metrics reliably detect *L. goodei* mate preference. The well-documented effects of reinforcement in the *Lucania* system led us to consider assays or metrics to be reliable if they detected conspecific mate preference for *L. goodei*, and if the strength of conspecific preference roughly agreed with the estimates of other assays or metrics within this study and with the estimates of strength of preference from the literature. The second goal was to use multiple assays and metrics to determine if *L. goodei* have native mate preferences consistent with cascade reinforcement. The third goal was to use multiple assays and metrics to determine if mate preference (either at the between- or within-species level) varies between sexes. Using multiple assays and metrics for the final 2 goals allowed us to corroborate our mate preference

findings. We measured conspecific and native mate preference for males and females using 3 different assays and 3 different metrics of preference.

METHODS

Collection and care

This experiment required us to use 3 populations: 1) a focal population of *L. goodei*, to measure mate preference, 2) a heterospecific population of a closely related species to be used as stimulus mates, and 3) a second, distinct population of *L. goodei*, to be used as foreign stimulus mates. We collected the focal population of *L. goodei* from the Lower Bridge of the Wakulla River in northern Florida (Wakulla County, Florida). Previous studies have repeatedly shown high levels of conspecific preference in this population (Fuller 2008; Gregorio et al. 2012; Kozak et al. 2015). We also collected heterospecific stimulus mates (*L. parva*) from this same population. Finally, we collected foreign *L. goodei* stimulus mates from Blue Springs, Florida (Gilchrist County, Florida), a site that occurs in a separate drainage (Suwanee) and that differs in mtDNA sequence (Murphy, unpublished data). We used dip nets and seines to collect males and females from each population during the summers of 2015 and 2016. Fish were transported in coolers back to the University of Illinois Urbana – Champaign, where they were housed in stock tanks in a greenhouse. Fish were exposed to natural light cycles and fed a diet of brine shrimp and blood worms daily.

Administration of assays

From our collected fish, we randomly selected 10 male and 10 female *L. goodei* individuals from the Lower Bridge population to be used as focal individuals. Each focal individual experienced 2 rounds of testing. The first round sought to measure conspecific mate preference, whereas the second round sought to measure native versus foreign mate preference. Each round of testing occurred over a period of 9 days. Identical methods were used with the exception of the identity of the stimulus mates (conspecific vs. heterospecific; native vs. foreign). On day 1, focal individuals took part in a dichotomous choice assay immediately followed by an audience assay. On day 2, both assays were administered a second time. Once assays were completed on day 2, one of the two stimulus mates (from the dichotomous and audience assays) was randomly assigned to remain with the focal individual for the no-choice assay and the other stimulus mate was removed. The pair was given 24 h to acclimate to the tank and then a no-choice assay began. The no-choice assays lasted for 7 days, and eggs were collected and counted from each pair daily. At the end of this first round of testing, stimulus mates were removed and focal individuals were given a rest period of several days. After the rest period, the above process was repeated using native and foreign stimulus mates.

Dichotomous choice assay

This assay involved a free swimming focal individual (either a male or female *L. goodei*) choosing between 2 caged stimulus mates. Focal individuals were placed into 38-L tanks at least 24 h prior to the experiment. Each tank contained a spawning mop (~20 pieces of green yarn tied together and weighted to sit on the bottom of the tank). The mop served as a spawning substrate and as a place to hide for the focal individual. Immediately before the start of the experiment, we placed 2 mesh cages 1.5 cm below the tank waterline in the front 2 corners of the tank. We randomly assigned and placed stimulus mates into the cages and gave them 10 min to acclimate to the new environment.

During the acclimation period, the cages rested near the front of the tank just below water level as to not disturb the focal individual. After the acclimation period, we gently lowered the cages to rest on the bottom of the tank, about 2.5 cm away from the corners. Moving the cages gave the focal individual freedom to approach stimulus mates from above and from all 4 sides. Once stimulus mate cages rested on the bottom of the tank, we began the assay which lasted for 10 min. During the assay, we recorded the amount of time focal individuals spent within 1 body length of each stimulus mate and the number of courting bouts performed by males. We measured courting bouts for both the focal males (i.e., males choosing among females) and the stimulus males (i.e., males being chosen by females). This assay was repeated on days 1 and 2.

From these data, we calculated conspecific and native preference for both males and females using both time associated with each stimulus mate and courting bouts. For males, we recorded the number of courting bouts directed towards each stimulus mate. For females, we recorded the number of courting bouts received from each stimulus mate. Although courtship is performed exclusively by males, using courtship as a proxy for female preference is not uncommon (Wagner 1998). In nature, once males start courting females, the females may either stay (and continue to be courted) or they may leave. Here, we used number of courting bouts as a proxy for time spent in courtship with a stimulus mate and therefore use it as a measurement of both male and female preferences.

We first calculated time-preference as the signed difference in association time between conspecifics and heterospecifics (i.e., time spent with conspecifics minus time spent with heterospecifics) for each focal individual for each day. Likewise, we calculated courtship-preference as the difference in courtship bouts (either given or received) between conspecifics and heterospecifics for each focal individual for each day. Assays where focal individuals spent no time with stimulus mates or where no courting was performed were not considered in analysis. The adjusted sample sizes reflect this in Tables 1 and 2. We next asked whether preferences differed between days 1 and 2 and as an effect of the order of treatments. We found no differences (see below; Supplementary Tables S1–S3), and we subsequently summed the association times and the courtship bouts and then calculated their signed differences. The same metrics were calculated for native versus foreign preferences. Positive values indicate preference for conspecific (or native) mates, whereas negative values indicate preference for heterospecific (or foreign) mates. We tested whether preferences differed from a null expectation of zero (no preference) using a 1-sample *t*-test.

Audience assay

Audience assays provided the focal individual the choice between a restrained mate and a free-swimming mate. The audience assays allowed males and females to interact in a natural fashion while preventing competition among stimulus mates. By having an alternate, caged stimulus mate present, the audience assay also provided a potential comparison to the freely available mate. Such a comparison is absent in the no-choice assays.

At the end of the dichotomous choice assays, 2 restrained stimulus mates rested on the bottom of the tank in their respective cages. Following a dichotomous choice assay, a stimulus mate was randomly

Table 1

Reproductive Isolation (RI) values representing conspecific (positive numbers) or heterospecific (negative numbers) mate preferences for male and female *L. goodei*

Assay	Sex	Metric	N	RI	CI
Dichotomous choice assay	Male	Time	10	0.71	0.49, 0.93
		Courting	10	0.81	0.63, 0.99
	Female	Time	9	0.02	−0.31, 0.34
		Courting	7	1	1, 1
Audience assay	Male	Time	10	0.51	0.11, 0.91
		Courting	10	0.55	0.11, 0.98
	Female	Time	10	0.5	0.15, 0.84
		Courting	9	0.95	0.87, 1.04
No-choice assay	Male	Eggs	10	0.55	0.37, 0.74
	Female	Eggs	10	1	1, 1

RI values were calculated for 3 metrics and 3 assays.
CI = 95% confidence intervals. CIs were calculated via bootstrapping for the no-choice assays.

Table 2

Reproductive Isolation (RI) values representing native (positive numbers) or foreign (negative numbers) mate preference for male and female *L. goodei*

Assay	Sex	Metric	N	RI	CI
Dichotomous choice assay	Male	Time	9	0.04	−0.36, 0.45
		Courting	7	−0.31	−0.8, 0.18
	Female	Time	10	0.01	−0.35, 0.37
		Courting	7	0.66	0.22, 1.09
Audience assay	Males	Time	9	0.05	−0.48, 0.58
		Courting	9	0.13	−0.39, 0.65
	Female	Time	10	0.64	0.39, 0.89
		Courting	9	0.89	0.63, 1.15
No-choice assay	Male	Eggs	10	−0.18	−0.57, 0.053
	Female	Eggs	10	0.30	−0.21, 0.93

RI values were calculated for 3 metrics and 3 assays.
CI = 95% confidence intervals. CIs were calculated via bootstrapping for the no-choice assays.

chosen to be released. Once free, the empty cage was removed from the tank and the audience assay began. Assays lasted for 10 min. During the assay, we recorded the amount of time the focal individual spent within 1 body length of each stimulus mate (either caged or free) and the number of courting bouts performed by males (either as stimulus mates or as focal individuals). We repeated this assay on day 2, but reversed which mate was released so that the caged stimulus mate from day 1 was freed, and the free stimulus mate from day 1 was caged.

The statistical methods used to measure preference for conspecific and native mates in the audience assay were identical to those used in the dichotomous choice assays. We summed the amount of time spent with conspecifics and with heterospecifics (regardless of whether they were free or caged) across the 2 days and calculated their signed difference. Likewise, we summed the amount of courtship given (males) or received (females) across the 2 days and calculated the difference between conspecific and heterospecific mates. Assays where focal individuals spent no time with stimulus mates or where no courting was performed were not considered in analysis. Qualitatively identical results were obtained when we compared preference for free mates (i.e., time spent with free conspecific vs. free heterospecific on separate days) or when we compared preference for caged mates (i.e., time spent with caged conspecific vs. caged heterospecific on separate days). We used the same statistical methods to measure preferences for native versus foreign mates. We tested whether preferences differed from a null expectation of zero using a 1-sample *t*-test.

No-choice assay

No-choice assays compared the total number of eggs produced between conspecific and heterospecific mate pairs. No-choice assays did not provide the focal individuals a choice between mates and instead paired them with a single mate for 7 days. For conspecific preference trials, we randomly assigned either a conspecific or heterospecific stimulus mate to each focal fish. Likewise, for native versus foreign preference trials, we randomly assigned either a native or foreign stimulus mate to each focal fish. The focal individual was paired with one of the two randomly chosen stimulus mates. The mate was placed into the tank along with 2 floating and 2 bottom yarn mops. The mops provided spawning substrate for the fish. The floating mops were attached to Styrofoam balls to allow them to float at the top of the tank, whereas the bottom mops were attached to PVC pipe and laid on the bottom of the tank. After the stimulus mate was added to the tank, the pair was given 24 h to acclimate. After the acclimation period, we collected and counted eggs from the mops each morning for 7 days. At the end of the assay period, stimulus mates were removed from the tank and returned to stock tanks.

Here, we simply measured preference as the total number of eggs spawned. For each sex, we tested whether there was a difference between the number of eggs produced with conspecifics versus heterospecifics or between native versus foreign mates using either a 2-sample *t*-test or a Kruskal–Wallis nonparametric test. We used the Kruskal–Wallis test when the assumptions of the

parametric *t*-test were violated. This happened when zero eggs were produced between *L. goodei* females and heterospecific males. All statistical tests were 2-tailed. All analyses were performed in *R* (version 1.0.136).

Order effects

The same stimulus and focal animals were used on days 1 and 2 for the dichotomous and audience assays. The focal individual and 1 of the 2 stimulus mates were used in the no-choice assay. This type of repeated testing was necessary so that we could control for the reproductive state and time of year of males and females. Still, the experimental design raises the possibility that the order of the administration of treatments may have affected the preference measures. We can imagine a variety of scenarios that could alter preferences. First, preferences could become stronger over time because individuals have more information about their potential mates. Conversely, preferences might become weaker over time due to experimental fatigue. We tested for these effects in the dichotomous and audience assays by comparing preference values between days 1 and 2 with linear and generalized linear models (Supplementary Tables S1–S3). We found no support for such effects.

Another possibility is that the administration of the audience assay on day 1 altered the preferences on day 2. Focal individuals had free access to one of the two stimulus mates (conspecific vs. heterospecific or native vs. foreign) during the assay. Using linear and generalized linear models, we asked whether the order of the presentation of the stimulus mates altered the preference scores between days 1 and 2 for the dichotomous and audience assays. We also asked whether the order of the audience assay affected the number of eggs laid in the no-choice assay (Supplementary Tables S4–S7). Again, we found no significant differences due to order. These results are summarized in Supplementary Material.

Quantifying reproductive isolation

For all 3 assays, we calculated RI for each metric.
To calculate RI, we used Stalker's (1942) equation for RI:

$$\frac{(Conspecific\ metric - Heterospecific\ metric)}{(Conspecific\ metric + Heterospecific\ metric)}$$

This equation creates a relative measure of preference, allowing for the direct comparison of the different metrics. The measure of RI ranges from −1 to 1, with negative numbers representing heterospecific (or foreign) preferences and positive numbers representing conspecific (or native) preferences. For the time and courting metrics for both the dichotomous choice and audience assays, we calculated RI for each individual and then calculated the 95% confidence intervals around the population mean. However, we could not directly apply this formula to individuals from no-choice data since focal individuals were only exposed to 1 mate. Instead, we used a bootstrap resampling method with replacement to determine mean RI and 95% confidence interval for each sex (10,000 replicates). The population level RI was calculated as the scaled difference in the number of eggs laid with conspecific (or native) versus heterospecific (or foreign) mates. We calculated RI for each of the 10,000 replicates and used these to calculate the 95% confidence interval.

Finally, we used Stalker's formula to calculate RI values for *L. goodei* using data from previous studies (Table 3). We used these RI values as a baseline to evaluate our newly calculated RI values. In addition to the criteria that reliable metrics and assays would detect a statistically significant preference for conspecific mates and that they would roughly agree with one another, we included the additional criterion that reliable metrics and assays would have RI values consistent with those from previous studies.

Table 3

Reproductive Isolation (RI) values calculated using data from previous reinforcement studies in *L. goodei*

Study	Assay type	Sex	Metric	RI
Kozak et al. (2015)	Dichotomous choice assay	Male	Time	0.31
		Female	Time	0.44
Fuller et al. (2007)	No-choice assay	Male	Eggs	0.44
		Female	Eggs	0.47

The raw data associated with this study have been submitted to Dryad (10.5061/dryad.1n1b75b).

RESULTS

Determining reliable assays and metrics

The first goal was to determine which assays and metrics reliably detected conspecific mate preference for *L. goodei*. To be considered reliable, metrics must be consistent with one another and with previous studies' estimates. Male conspecific mate preference was found in all 3 assays (Figure 1, Table 1). Males spent significantly more time with conspecific mates and courted them more often during dichotomous choice assays (Figure 1a,b). The RI values for dichotomous choice metrics were also consistent with one another (RI-time = 0.71, RI-courting = 0.81), and their 95% confidence intervals overlap (Table 1). During audience assays, males courted conspecific mates significantly more often, but only spent marginally more time with them (Figure 1a,b). RI values for both metrics exceeded RI values calculated from previous studies, but were lower than dichotomous choice RI values (Table 1). Finally, no-choice assays also detected male mate preference as assays with conspecific pairs produced significantly more eggs than assays with heterospecific pairs (Figure 1c, Table 1). Furthermore, the RI value for no-choice assays was similar to those found in the audience assay (~0.5) (Table 1).

Assays and metrics varied much more in their ability to detect female mate preference. During dichotomous choice assays, females did not prefer to spend time with one mate over the other (Figure 1a). Relatively few courting bouts occurred in the female dichotomous choice assays, but the courtship that females did receive was solely from conspecifics (Figure 1b). The value for RI clearly varied between the 2 metrics. The RI for the time metric severely underestimated female mate preference (RI-time = 0.02), whereas the RI for the courting metric reported significant conspecific preference (RI-courting = 1.00) (Table 1). Audience assays did detect female conspecific mate preference. Females spent significantly more time with and were courted more often by conspecific mates (Figure 1a,b). However, the RI value for courting was nearly twice as high as the RI value for time in the audience assays (Table 1). Finally, the no-choice assay also detected conspecific mate preference. Females produced significantly more eggs with conspecific partners than with heterospecific partners (Figure 1c). In fact, zero eggs were produced during the entirety of the assay period when female *L. goodei* were paired with heterospecific mates. Since females produced zero eggs with heterospecifics, RI for the no-choice assays was very high (RI-eggs = 1.00) (Table 1). The RI

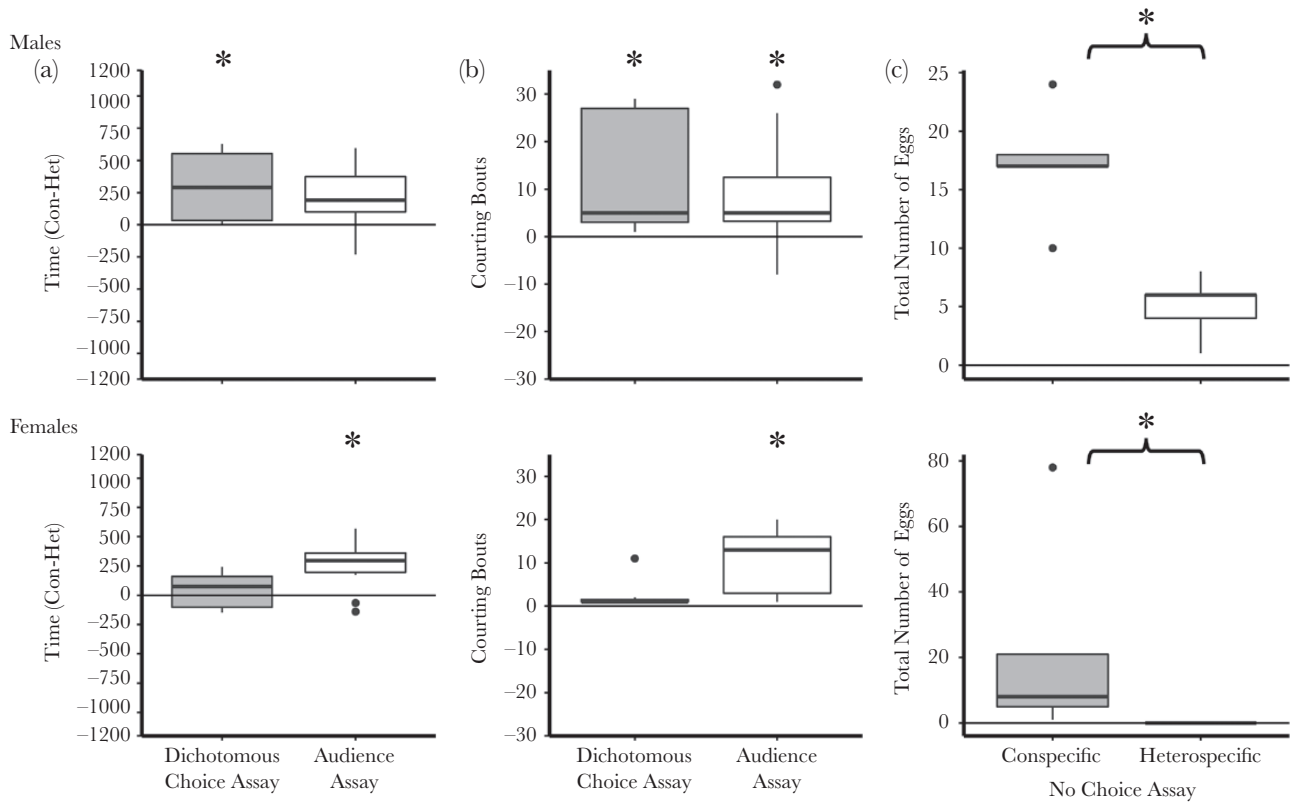

Figure 1
Boxplots indicating male (top) and female (bottom) mate preference for conspecific (positive numbers) or heterospecific (negative numbers) mates measured with (a) the time metric, (b) the courting bouts metric, or (c) the total number of eggs produced metric. Asterisks (*) represent statistically significant differences from zero for the (a) time metric and the (b) courting bouts metric and significant differences in (c) number of eggs laid between conspecifics and heterospecifics in the no-choice assay.

values for courting in the dichotomous choice and audience assays and the RI values for the no-choice assays were all similar to one another.

Native versus foreign mate preference

The second goal of this experiment was to use several metrics and assays to determine if Lower Bridge *L. goodei* preferred native or foreign mates. We first looked for mate preference in Lower Bridge males, but found that none of our assays or metrics detected a significant preference for either mate (Figure 2). The RI calculations also reflected this lack of preference (Table 2). The RI values of 2 metrics, however, did stand out. The courting metric from the dichotomous choice assays and the total number of eggs produced metrics from the no-choice assays detected the strongest mate preference in males (RI-courting = −0.31 and RI-eggs = −0.18) (Table 2). Remarkably, both RI calculations indicate that males may prefer foreign mates.

In contrast, female preference for native mates was detected in some assays. Females were courted significantly more often by native mates during dichotomous choice assays (Figure 2b). They also spent more time with and were courted more often by native mates during audience assays (Figure 2a,b). RI values for each of these metrics also matched this pattern (0.66, 0.64, and 0.89, respectively) (Table 2). The RI value for the no-choice assay was 0.30 and was not statistically different from zero, but there was a large outlier (Figure 2c). Removal of this data point results in a significant preference for native males. Overall, Lower Bridge females, but not males, appear to have a within-species mate preference for native mates.

Male and female comparison

Although Lower Bridge males and females both preferred conspecific mates, female preference was stronger than male preference. Female mate preference was stronger for 1) the courting metric in the dichotomous choice assay, 2) the courting metric in audience assays, and 3) the total number of eggs produced metric in no-choice assays (Table 1). RI values for these metrics were not only higher than male RI values, but they were also similar to one another (Table 1). The time metric was the most inconsistent mate preference measurement between males and females. In the dichotomous choice assay, male RI was significantly higher than female RI, but the audience assay detected no difference between the sexes (Table 1).

Native mate preference also differed between male and female *L. goodei*. Native mate preference was present in *L. goodei* females, but not males. For females, both time and courtship in the audience assays and courtship received in the dichotomous choice assay revealed high RI values with overlapping 95% confidence limits (Table 2). The time metric in the dichotomous choice assay, again, underestimated RI values for females. For males, all metrics showed no significant preference for native females, and, if anything, revealed a slight preference for foreign mates (Table 2).

DISCUSSION

The first goal of this study was to determine which metrics or assays reliably detected *L. goodei* mate preference. Ultimately, we identified 2 metrics and assays that met our reliability criteria. First, the courting metric from audience assays detected that both males

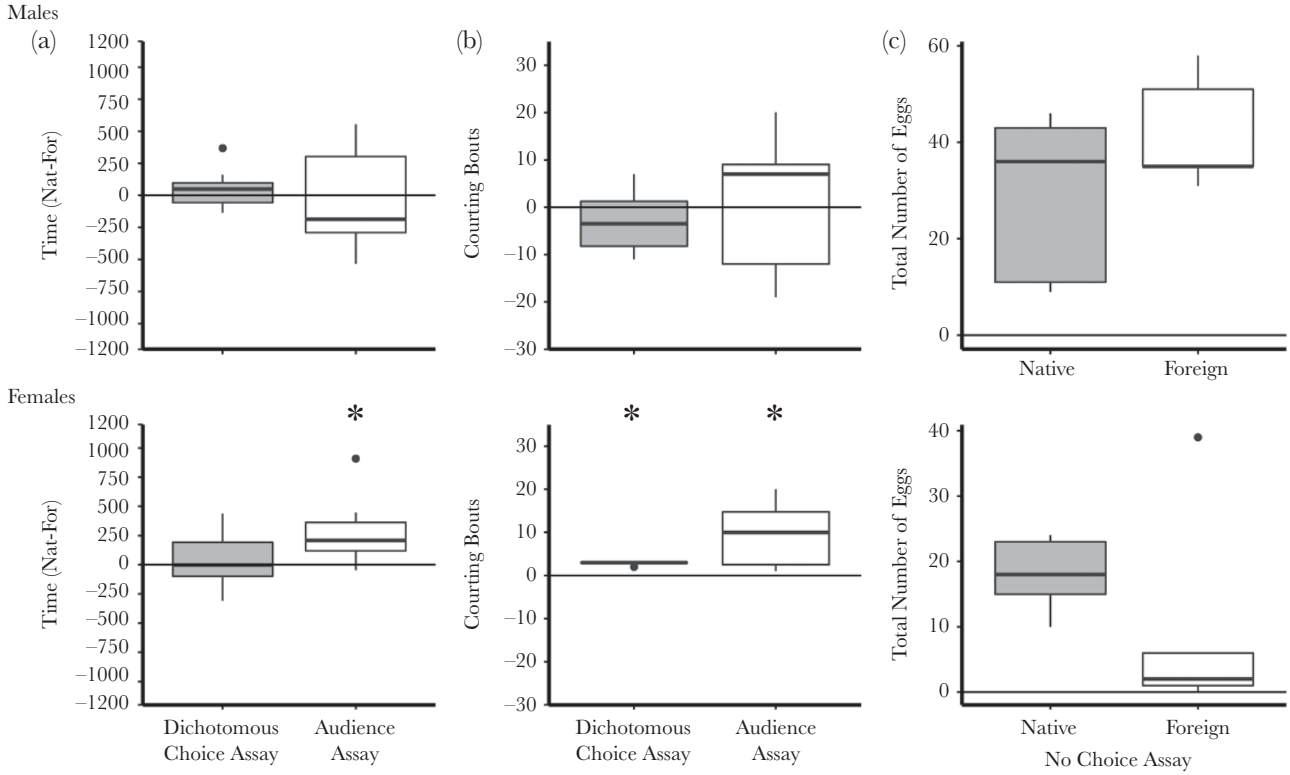

Figure 2
Boxplots indicating male (top) and female (bottom) mate preferences for native (positive numbers) or foreign (negative numbers) mates measured with (a) the time metric, (b) the courting bouts metric, or (c) the total number of eggs produced metric. Asterisks (*) represent statistically significant preferences from zero for the (a) time metric and (b) the courting bouts metric.

and females courted or were courted by conspecific mates more often than heterospecific mates (fulfilling criteria 1; Figure 1b). The RI values for both males and females were also roughly consistent with those found using other assays and metrics in this study and with previous estimates in the literature (fulfilling criteria 2; Tables 1 and 3). We note that the audience assay is somewhat novel. Although dichotomous choice and no-choice assays have been used for decades (reviewed in Andersson (1994) and Rosenthal (2017)), the audience assay has not been used as much. The advantage of the audience assay is that it allows for more natural interactions between males and females while both 1) preventing overt competition between members of the same sex and 2) reminding the focal individual of the presence of other potential mates in the population.

The second reliable metric was the total number of eggs produced from the no-choice assay. Conspecific pairings produced significantly more eggs than heterospecific pairings for both males and females (fulfilling criteria 1; Figure 1c). Additionally, the RI values for both males and females were roughly consistent with those found using other assays and metrics in this study and with previous estimates in the literature (fulfilling criteria 2; Tables 1 and 3). The RI values from the courting metric and the total number of eggs produced metric were also consistent with one another for both males and females (fulfilling criteria 2; Table 1). Undoubtedly, an increased sample size may have resulted in narrower confidence limits, which may have rendered some of our RI values significantly different from zero (i.e., number of of eggs laid by females with native vs. foreign males). Still, our study found that behaviors linked to courting produced robust estimates of RI that were concordant with previous studies. These results were somewhat surprising. Previous studies concluded that association time predicted mate preference more reliably than metrics specifically measuring mating behaviors, such as gliding (Cummings and Mollaghan 2006). Cummings and Mollaghan (2006) argued that association time was more reliable because individuals could associate with prospective mates regardless of breeding condition. Although this may be true for organisms who only approach or spend time with desired mates, it does not fit with the *L. goodei* mating ritual. Male *L. goodei* establish territories, whereas female *L. goodei* sequentially visit said territories (Arndt 1971; Fuller 2001; McGhee et al. 2007). A female may visit, and thus spend time with, a male who she ultimately does not mate with (Fuller 2001). Sexes may also associate for reasons unrelated to reproduction. For example, some species may associate with other fish to reduce the risk of predation (Brock and Riffenburgh 1960) or may associate with other fish in schools to increase the hydrodynamics of cruising (Pitcher et al. 1985). Also, association time may not be indicative of mating preference if animals are not ready to mate. In many species, females are receptive to mating only during certain physiological states (e.g., after ovulation and after parturition in live bearers) (Liley and Stacey 1983; Constantz 1989; Breder and Rosin). The advantage of metrics that are associated with mating (i.e., courting bouts and number of eggs laid) is that they are unambiguously related to mating.

The drawback of using behaviors associated with reproduction is that it can be difficult to determine the contributions of males and females to preference. No-choice assays rely on the number of eggs

spawned as a metric of preference. Yet spawning requires the participation of both males and females. When a large number of eggs are spawned, is this because the female was attracted to the male or because the male was willing to court and stimulate the female? Likewise, the amount of courtship received by females requires that 1) females are in close proximity to males, 2) males court females, and 3) females do not swim away. When no courtship is received, is this due to the fact that a given female did not like a male or because the male found the female unattractive and never courted or both? These problems seem obvious for traits such as number of eggs produced and the amount of courtship received, but preferences based on association time suffer from the same problem when live animals are used as stimuli. Do individuals genuinely prefer 1 type of stimulus mate over another or are the stimuli reacting differently to the focal individual, which influences preference? Both statistical and experimental approaches can be taken to determine the likely roles of males and females on preference. Yet, male and female preferences do not simply mirror one another. Instead, our study found unique patterns of preference for the different sexes in different settings (i.e., conspecific vs. heterospecific; native vs. foreign). Other work suggests that there are contributions of both sexes to preference and RI that can be resolved with experimental and statistical approaches (St. John 2017).

One caveat to this study is that we administered assays in a set order: the dichotomous choice assay was always followed by the audience assay (repeated over 2 days), and no-choice assays were always performed after the choice assays. This order was maintained in an effort to control the reproductive status of the female. Our fear was that females might not be receptive to males if they spawned all their eggs (Liley and Stacey 1983). Hence, the dichotomous choice trials preceded the audience assay (where the female had limited access to the male), which proceeded the no-choice trials where the animals could spawn for multiple days. Also, we feared that conducting the no-choice assay prior to the dichotomous choice and audience assays might cause fish to prefer new, novel mates (Eakley and Houde 2004; Hampton et al. 2009). Still, the experimental design may have introduced confounding order effects. First, it is possible that either focal individuals or stimulus mates may have become fatigued over the course of the experimental period, which would have reduced preference levels over time. Ideally, all assays would be administered in a random order to help control for this problem. However, in the wild, killifish have been observed mating continuously throughout the day and observed mating multiple times in a single 30-min observation making experimental fatigue less likely (Fuller 2001). Furthermore, although dichotomous choice and audience assays were always performed back-to-back, the total amount of time that focal or stimulus mates spent in trials for a single day was no more than 30 min, after which they were returned to isolation tanks until the next day. Another possibility is that focal individuals would gain more information about potential mates over time, which would increase preference levels over time. We found no difference between the 2 days nor did we find differences that were attributable to the order of stimulus mate access in the audience assay, indicating that the increased assessment time did not significantly affect detection of mate preference.

Rosenthal (2017) gives a thorough review of the costs and benefits of different experimental designs for measuring mating preferences (see also Houde (1997) and Wagner (1998)). Studies of mate choice need to consider whether to use 1) actual live mating events versus behaviors presumed to indicate preference, 2) free-spawning assays where animals can fully interact versus various designs for restraining stimuli, 3) behaviors associated with different stages of the mating ritual, and 4) live animals versus experimentally manipulated animals versus synthetic animals, which allow for different levels of signal manipulation. Clearly, our understanding of mating preferences will be best when we take multiple approaches to measuring preference and understanding its manifestation in nature. The current study reflects this sentiment as we have combined multiple estimates of preference with previous estimates that offer a robust picture on preferences for conspecific versus heterospecific and native versus foreign fish.

The second goal of this study was to determine whether cascade reinforcement might be present. Specifically, we sought to determine whether *L. goodei* from the Lower Bridge population preferred native or foreign mates. We found that female *L. goodei* preferred native mates, whereas male *L. goodei* showed no mate preference (Table 2). Native mate preference in female *L. goodei* is not only consistent with the prediction that reinforcement can cascade into within-species preferences, but also mirrors results from previous studies. Kozak et al. (2015) found that female *L. parva* exhibited native mate preference, whereas male *L. parva* showed no preference for native or foreign mates. Although this finding is a first step in determining if cascade reinforcement is occurring in *L. goodei*, it does not confirm cascade reinforcement per se. The signature of cascade reinforcement is strengthened native mate preferences in sympatry compared with allopatry (Ortiz-Barrientos et al. 2009). Future studies comparing female native mate preferences in sympatry and allopatry are needed to confirm cascade reinforcement in *L. goodei*.

Finally, the last goal of this study was to determine whether the 2 sexes differed in conspecific and native mating preferences. We found that *L. goodei* females had stronger mate preference than males. Although both sexes exhibited a significant preference for conspecific mates, RI values for females were consistently close to 1 while RI values for males were closer to ~0.55 (Table 1). Female *L. goodei* were also the only sex to exhibit significant native mate preference (Table 2). The stronger mate preference of female *L. goodei* supports the hypothesis that larger energetic investment in reproduction by females drives asymmetric prezygotic isolation in the *Lucania* system. This was unexpected, because hybrid offspring with *L. goodei* fathers (and *L. parva* mothers) suffer higher fitness costs than hybrid offspring with *L. goodei* mothers (and *L. parva* fathers) (Fuller 2008). The amount of energy invested by females in reproduction (and potentially lost during heterospecific matings) outweighs the costs that males incur from unfit offspring. Differential investment in reproduction between sexes is well-documented. This study provides early support for how reinforcement subsequently acts on the differences between sexes (Livingstone 1974; Wigby and Chapman 2005; Hayward and Gillooly 2011; Rankin et al. 2011).

In conclusion, we found that *L. goodei* mate preference was best detected using behaviors associated with mating, such as egg production or courting. Using these metrics, we supported the finding of conspecific mate preference in sympatric *L. goodei*. We added further to this, by documenting differences in the strength of mate preference between sexes, where female *L. goodei* had much stronger conspecific mate preference than males. Not only did females have stronger conspecific mate preference, but they were also the only sex to exhibit preferences for males from their own, native populations over foreign populations. Our findings support the hypothesis that differential investment in reproductive events can affect the formation of prezygotic isolation via reinforcement and its cascading effects.

SUPPLEMENTARY MATERIAL

Supplementary data are available at *Behavioral Ecology* online.

ACKNOWLEDGMENTS

This work was funded by the National Science Foundation (DEB 0953716), the UI Department of Animal Biology Odum-Kendeigh Fund, and the UI School of Integrative Biology Harley and Francis Clark Fund. M.E. St. John was supported by a National Institutes of Health Science Education Partnership Award (R25 OD020203) to B. Hug and R.C.F. L.D. Mitchem, J. Knox, and Z. Osterholz helped collect fish. L.D. Mitchem, C.-H. Chang, R. Moran, A. Bell, S. Berlocher, and 2 anonymous reviewers provided comments that improved the manuscript. These experiments were approved by Illinois Institutional Animal Care and Use Committee (#15246).

Data accessibility: Analyses reported in this article can be reproduced using the data provided by St. John and Fuller (2018).

Handling Editor: Ulrika Candolin

REFERENCES

Andersson MB. 1994. Sexual selection. Princeton: Princeton University Press.

Arndt RGE. 1971. Ecology and behavior of the cyprinodont fishes Adinia xenica, Lucania parva, Lucania goodei, and Leptolucania ommata. [PhD dissertation]. [Ithaca, NY]: Cornell University.

Berdan EL, Fuller RC. 2012. A test for environmental effects on behavioral isolation in two species of killifish. Evolution. 66:3224–3237.

Blair WF. 1974. Character displacement in frogs. Am Zool. 14: 1119–1125.

Bolnick DI, Turelli M, López-Fernández H, Wainwright PC, Near TJ. 2008. Accelerated mitochondrial evolution and "Darwin's corollary": asymmetric viability of reciprocal F1 hybrids in Centrarchid fishes. Genetics. 178:1037–1048.

Breder CM, Rosin DE. Modes of reproduction in fishes. Garden City (NY): The Natural History Press.

Brock VE, Riffenburgh RH. 1960. Fish schooling: a possible factor in reducing predation. ICES J Mar Sci. 25:307–317.

Charmantier A, Brommer JE, Nussey DH. 2014. The quantitative genetics of senescence in wild animals. Quant Genet wild. :68–83.

Constantz GD. 1989. Reproductive biology of poeciliid fishes. Ecol Evol Livebearing Fishes. :33–50.

Constantz GD. 1989. Reproductive biology of Poeciliid fishes. In: Meffe GK, Snelson FF, editors. Ecology and evolution of livebearing fishes (Poeciliidae). New York: Prentice Hall. p. 33–50.

Cotton S, Small J, Pomiankowski A. 2006. Sexual selection and condition-dependent mate preferences. Curr Biol. 16:R755–R765.

Coyne JA, Orr HA. 2004. Speciation. Sunderland (MA): Sinauer Assoc.

Cummings M, Mollaghan D. 2006. Repeatability and consistency of female preference behaviours in a northern swordtail, *Xiphophorus nigrensis*. Anim Behav. 72:217–224.

Dougherty LR, Shuker DM. 2015. The effect of experimental design on the measurement of mate choice: a meta-analysis. Behav Ecol. 26:311–319.

Eakley AL, Houde AE. 2004. Possible role of female discrimination against 'redundant' males in the evolution of colour pattern polymorphism in guppies. Proc Biol Sci. 271 (Suppl 5):S299–S301.

Fuller RC. 2001. Patterns in male breeding behaviors in the bluefin killifish, *Lucania goodei*: a field study (Cyprinodontoformes: Fundulidae). Copeia. 2001:823–828.

Fuller RC. 2008. Genetic incompatibilities in killifish and the role of environment. Evolution. 62:3056–3068.

Fuller RC. 2016. Reconciling concepts, theory, and empirical patterns surrounding cascade reinforcement. Curr Zool. 62:131–134.

Fuller RC, McGhee KE, Schrader M. 2007. Speciation in killifish and the role of salt tolerance. J Evol Biol. 20:1962–1975.

Gregorio O, Berdan EL, Kozak GM, Fuller RC. 2012. Reinforcement of male mate preferences in sympatric killifish species *Lucania goodei* and *Lucania parva*. Behav Ecol Sociobiol. 66:1429–1436.

Hampton KJ, Hughes KA, Houde AE. 2009. The allure of the distinctive: reduced sexual responsiveness of female guppies to "redundant" male colour patterns. Ethology. 115:475–481.

Hayward A, Gillooly JF. 2011. The cost of sex: quantifying energetic investment in gamete production by males and females. PLoS ONE. 6:e16557.

Hoffmann AA, Turelli M. 1997. Cytoplasmic incompatibility in insects. Influ Passengers Inherit Microorg Arthropod Reprod. 140:42–80.

Hopkins R, Guerrero RF, Rausher MD, Kirkpatrick M. 2014. Strong reinforcing selection in a Texas wildflower. Curr. Biol. 24:1995–1999.

Hoskin CJ, Higgie M. 2010. Speciation via species interactions: the divergence of mating traits within species. Ecol Lett. 13:409–420.

Houde A. 1997. Sex, color, and mate choice in guppies. Princeton University Press.

Hunt J, Brooks R, Jennions MD. 2005. Female mate choice as a condition-dependent life-history trait. Am Nat. 166:79–92.

Jaenike J, Dyer KA, Cornish C, Minhas MS. 2006. Asymmetrical reinforcement and Wolbachia infection in Drosophila. PLoS Biol. 4:e325.

Kelly JK, Noor MA. 1996. Speciation by reinforcement: a model derived from studies of Drosophila. Genetics. 143:1485–1497.

Kokko H, Jennions MD. 2015. Describing mate choice in a biased world: comments on Edward and Dougherty & Shuker. Behav Ecol. 26:320–321.

Kozak GM, Roland G, Rankhorn C, Falater A, Berdan EL, Fuller RC. 2015. Behavioral isolation due to cascade reinforcement in Lucania killifish. Am Nat. 185:491–506.

Lemmon EM. 2009. Diversification of conspecific signals in sympatry: geographic overlap drives multidimensional reproductive character displacement in frogs. Evolution. 63:1155–1170.

Livingstone FB. 1974. Sexual selection and the descent of man 1871–1971. Am J Phys Anthropol. 40:293–295.

McGhee KE, Fuller RC, Travis J. 2007. Male competition and female choice interact to determine mating success in the bluefin killifish. Behav Ecol. 18:822–830.

Moran RL, Fuller RC. 2018. Agonistic character displacement of genetically based male colour patterns across darters. Proc R Soc B Biol Sci. 285:20181248.

Moran RL, Zhou M, Catchen JM, Fuller RC. 2017. Male and female contributions to behavioral isolation in darters as a function of genetic distance and color distance. Evolution. 71:2428–2444.

Nosil P. 2007. Divergent host plant adaptation and reproductive isolation between ecotypes of Timema cristinae walking sticks. Am Nat. 169:151–162.

Ortiz-Barrientos D, Grealy A, Nosil P. 2009. The genetics and ecology of reinforcement: implications for the evolution of prezygotic isolation in sympatry and beyond. Ann N Y Acad Sci. 1168:156–182.

Pfennig KS. 2016. Reinforcement as an initiator of population divergence and speciation. Curr Zool. 62:145–154.

Pfennig KS, Simovich MA. 2002. Differential selection to avoid hybridization in two toad species. Evolution. 56:1840–1848.

Pfennig KS, Rice AM. 2014. Reinforcement generates reproductive isolation between neighbouring conspecific populations of spadefoot toads. Proc R Soc B Biol Sci. 281:20140949.

Pitcher TJ, Magurran AE, Edwards JI. 1985. Schooling mackerel and herring choose neighbours of similar size. Mar Biol. 86:319–322.

Rankin DJ, Dieckmann U, Kokko H. 2011. Sexual conflict and the tragedy of the commons. Am Nat. 177:780–791.

Rosenthal GG. 2017. Mate choice: teh evolution of sexual decision making from microbes to humans. Princeton: Princeton University Press.

Rundle HD, Schluter D. 1998. Reinforcement of stickleback mate preferences: sympatry breeds contempt. Evolution. 52:200–208.

Ryan MJ, Taylor RC. 2015. Measures of mate choice: a comment on Dougherty & Shuker. Behav Ecol. 26:322–323.

Sætre G, Moum T, Bures S, Kral M, Adamjan M, Moreno J. 1997. A sexually selected character displacement in flycatchers reinforces premating isolation. Nature. 387:1995–1998.

St. John M, 2017. Reinforcement and cascade reinforcement in the Lucania system: the effects of experimental design, sex, and heterospecific pairings on mate preference. [M.Sc. dissertation] University of Illinois.

St. John M, Fuller R. 2018. Data from: the effects of experimental design on mating preferences and reproductive isolation in killifish. Dryad Digital Repository. https://doi.org/10.5061/dryad.1n1b75b

Liley NR, Stacey NE. 1983. 1 Hormones, pheromones, and reproductive behavior in fish. Vol. 9. Academic Press. 1–63.
Stalker HD. 1942. Sexual isolation studies in the species complex *Drosophila virilis*. Genetics. 27:238–257.
Wagner WE. 1998. Measuring female mating preferences. Anim Behav. 55:1029–1042.
Walker BW, Johnson RE. 1943. Hybridization in nature between species of American cyprinodont fishes. Contr Lab Vert Biol. 23: 1–21.
Wigby S, Chapman T. 2005. Sex peptide causes mating costs in female *Drosophila melanogaster*. Curr Biol. 15:316–321.
Wirtz P. 1999. Mother species-father species: unidirectional hybridization in animals with female choice. Anim Behav. 58:1–12.
Yukilevich R. 2012. Asymmetrical patterns of speciation uniquely support reinforcement in Drosophila. Evolution. 66:1430–1446.
Yukilevich R, True JR. 2006. Divergent outcomes of reinforcement speciation: the relative importance of assortative mating and migration modification. Am Nat. 167:638–654.