



Invited Review

Further mismeasures of animal contests: a new framework for assessment strategies

Kenneth James Chapin,^{a,b,○} Paulo Enrique Cardoso Peixoto,^{c,○} and Mark Briffa^{d,○}

^aDepartment of Ecology & Evolutionary Biology, University of California, Los Angeles, 612 Charles E. Young Drive East, Los Angeles, CA, USA, ^bDepartment of Ecology & Evolutionary Biology, The University of Arizona, P.O. Box 210088, Tucson, AZ 85721, USA, ^cDepartamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brasil, and ^dSchool of Biological & Marine Sciences, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK

Received 18 October 2018; revised 29 April 2019; editorial decision 1 May 2019; accepted 5 June 2019.

Competition for resources is a ubiquitous feature of life, and a central topic in behavioral ecology. Organisms use assessment strategies to resolve contests, which can be delineated into two broad categories by the information individuals use to make decisions: mutual assessment (MA) or self-assessment (SA). Most research hitherto has worked to bin a species into one of these categories. In this review, we discuss the limitations of this approach and provide solutions. We posit that assessment strategies do not need to be fixed within a species, individuals, or interactions, and that many organisms should adjust their assessment strategy as the environment, opponent, and opportunities for information gathering change. We show that assessment strategies are an individual-level characteristic, can vary within and between contests, and are not mutually exclusive. We argue that MA is the midpoint along a spectrum of self only and opponent only assessment. We discuss the effects of resource distribution, demographics, experience, information transfer, and ontogeny on assessment strategy evolution and behavior. We conclude by providing empirical guidelines and an example with a simulated dataset.

Key words: animal contests, assessment strategy, competition, fighting, mutual assessment, self-assessment, territoriality.

INTRODUCTION

When two individuals compete directly for resources (food, space, and mates), they must use some strategy to make contest decisions—assessment strategies. There has been a recent resurgence of interest in the study of assessment strategies (e.g., Bubak et al. 2016; Camerlink et al. 2016; Edmonds and Briffa 2016; Pajjmans and Wong 2017). This is in part due to the advances in Taylor and Elwood (2003), which described a new approach to test if opponents use information about each other to make contest decisions. While some studies have produced clear support that species use opponent information (e.g., Tibbetts et al. 2010; Schnell et al. 2015; Tedore and Johnsen 2015; Yasuda and Koga 2016) or do not (e.g., Prenter et al. 2006; Brandt and Swallow 2009; Constant et al. 2011), many have been less clear, oftentimes concluding a mix of strategies (e.g., Jennings 2005; Kelly 2006; Briffa 2008; Garcia et al. 2012; Peixoto and Benson 2012; Yasuda et al. 2012; Palaoro et al. 2014; Wofford et al. 2015; Edmonds and Briffa 2016). Thus, there is a need for both

extending and refining the original framework, both theoretical and empirical, used to understand contest assessment strategies. We posit that an individual-level approach to investigating assessment strategies will resolve these issues, and provide insight into the dynamic nature of assessment strategies in animal contests.

Investigations on the evolution of agonistic interactions have been ongoing for almost five decades (Arnott and Elwood 2008, 2009; Kokko 2013). Initially, resource contests were thought of as interactions between opponents where individuals make contest decisions by assessing their own and their opponents' resource holding potential (RHP, the absolute fighting ability of an individual; Parker 1974; Parker and Rubenstein 1981; Enquist and Leimar 1987; Enquist et al. 1990; Hardy & Briffa 2013). This type of assessment rule has been termed mutual assessment (MA). In MA, individuals gain information about the RHP of opponents. Thus, opponents can respond to relatively high-RHP opponents by fleeing before reaching extreme cost levels in the interaction. Models based on MA include the sequential assessment (Enquist et al. 1990) and asymmetric war of attrition models (Parker and Rubenstein 1981; Table 1).

Address correspondence to K.J. Chapin. E-mail: chapinkj@gmail.com.

Several researchers, however, have noted that not all animals assess their opponents during a contest (Maynard Smith 1974; Bishop and Cannings 1978; Mesterton-Gibbons et al. 1996; Taylor and Elwood 2003; Arnott and Elwood 2009; Elwood and Arnott 2012; Kokko 2013). Instead, animals use information only about their own fighting ability—not that of their opponent—to determine when to give up a contest (Taylor and Elwood 2003). A lack of opponent assessment can seem counterintuitive when agonistic interactions involve a series of stereotyped behaviors and, in some cases, injuries (Huxley 1966). That being said, both theoretical and empirical research shows that not all animals use opponent information in contests (Mesterton-Gibbons 1996; Prenter et al. 2006; Brandt and Swallow 2009; Constant et al. 2011). This can be imagined as follows: individuals might have a maximum display duration based on their own RHP. Individuals give up if an opponent continues to display or win if their opponent has given up. In this thought-example, individuals need not determine opponent RHP—only opponent presence. Self-assessment (SA) style models include war of attrition without assessment (Mesterton-Gibbons et al. 1996), the energetic war of attrition model (Payne and Pagel 1996, 1997; Table 1), and a special case of this model called the cumulative assessment model (CAM) (Payne 1998; but in terms of observed patterns, this model can yield similar predictions to an MA strategy, see below). Empirical research has identified self-assessment or a mix of self and MA strategies in diverse taxa (amphipods, Prenter et al. 2006; honeybees, Dietemann et al. 2008; house crickets, Briffa 2008; jumping spiders, Elias et al. 2008; shore crabs, Smallegange et al. 2007).

Over the last 15 years, the field of contest assessment has undergone revision in terms of how we empirically distinguish between mutual and self-assessment. Until 2003, researchers often used a correlation between the difference in RHP (usually body size or mass) and contest cost (most often measured as contest duration) to test the prediction that if rivals are using an MA strategy, then individuals with more similar RHP take longer to resolve contests (Enquist and Leimar 1983; Rosenberg and Enquist 1991; Morris et al. 1995; Mesterton-Gibbons et al. 1996; Hack 1997; Figure 1a). However, Taylor and Elwood (2003) showed that this correlation can occur, even if winner RHP does not correlate with contest duration (e.g., if individuals fight according to SA models; Taylor and Elwood 2003). To avoid this problem, Taylor and Elwood (2003) offered a solution: instead of analyzing RHP difference (a measure of “relative” RHP), correlate winner and loser RHP (i.e., “absolute” RHP) separately with contest duration (Taylor and Elwood 2003). While loser RHP should be positively associated with contest duration under both SA and MA, winner RHP should negatively correlate with contest duration only if opponents are using MA.

This approach (hereafter the “Taylor-Elwood approach”) has since been used extensively, but often with mixed or inconclusive results (Jennings 2005; Kelly 2006; Kemp 2006; Tibbetts et al. 2010; Tedore and Johnsen 2015; Ng et al. 2016; Yasuda and Koga 2016; Pinto et al. 2019). Further, the Taylor-Elwood approach did not consider other factors that may also affect contest dynamics, such as the importance that the disputed resource have to each rival. Although researchers have emphasized the importance of understanding how the value of a contested resource can vary among opponents and affect their decision rules (Parker and Rubenstein 1981; Enquist and Leimar 1987; Briffa and Elwood 2004; Arnott and Elwood 2008), this has not been included in a general empirical framework of assessment strategies. Perhaps, this can be resolved by an alternate approach that reveals individual-level differences in assessment strategies, relaxes several assumptions from game theoretic models, and incorporates resource value in assessment.

MODEL ASSUMPTIONS AND LIMITATIONS

Assessment strategies are treated as species- or population-level characteristics, but this is unlikely to be the case in natural populations (Taylor and Elwood 2003; Hsu et al. 2006; Prenter et al. 2006; Arnott and Elwood 2007, 2009; Camerlink et al. 2017). Instead, assessment strategies can vary 1) among individuals within populations, 2) within individuals across development, 3) within individuals during a contest, and 4) between environmental contexts. Indeed, there is empirical evidence of different assessment strategies among opponents in dyadic contests (Elwood et al. 1998; Briffa and Elwood 2004; Prenter et al. 2008; Arnott and Elwood 2009; Briffa and Lane 2017; Camerlink et al. 2017) and individual-level variation has been shown across many areas of animal behavior (Sih et al. 2004). Despite this, no research on how assessment strategies vary within populations has been published.

Contestants can change assessment strategies across stages of the contest (Hsu et al. 2006; Mesterton-Gibbons and Heap 2014). For example, the fiddler crab *Uca mjoebergi* first assesses opponent size to decide if it should engage in a fight (Morrell et al. 2005). If a fight initiates, contestants use SA to decide the interaction. This has been shown in several species, which oftentimes employ MA initially, followed by SA if the contest escalates to later stages (Moore et al. 2003; Morrell et al. 2005; Hsu et al. 2006; Stuart-Fox 2006; Arnott Elwood 2009; Lobregat et al. 2019). Despite this, many empirical studies assume that strategies are consistent throughout the contest.

Most studies, both empirical, and theoretical, assume that strategies do not vary between opponents. If this is the case for a given species, then the Taylor-Elwood approach only determines the

Table 1
Theoretical contest assessment strategies organized by information use

Strategy	S	O	R	Citation
War of attrition (WOA)	✓			Maynard Smith 1974
Asymmetric WOA	✓	✓		Parker and Rubenstein 1981
Energetic WOA (EWOA)	✓			Payne and Pagel 1996, 1997
EWOA without assessment (EWOA-WA)	✓			Mesterton-Gibbons et al. 1996
Sequential Assessment Model (SAM)	✓	✓		Enquist et al. 1990
Enquist-Leimar SAM	✓	✓	✓	Enquist and Leimar 1987
Cumulative Assessment Model (CAM)	✓			Payne 1998

S, O, and R, indicate self-assessment, opponent assessment, and resource assessment—the three primary information sources in animal contests.

assessment strategy of the loser but is often used to infer the assessment strategies of both opponents. This limitation occurs because the effect of the winner's strategy is not identified by the correlation of winner RHP with contest duration; these measures only identify the assessment strategy of the opponent that determines contest duration, which is intrinsically the contest loser. Basing our inferences about decision rules on the behavior of losers only is not problematic as long as we can assume that winners and losers use the same sources of information in similar ways. However, this assumption may not always be sound. In some examples, opponents employ different assessment strategies that use different sources of information (Arnott and Elwood 2009). In hermit crabs shell fights, for example, attackers have good information about the contested resource but poor information about the opponent's RHP, while the opposite is true for defenders (Briffa and Elwood 2004). Access to different types of information resulted in giving up driven by SA if attackers lose, but information about the opponent's RHP if defenders lose (Briffa and Elwood 2004). Even if opponents do not adopt distinct roles, winners and losers may have access to different

sources of information or be able to obtain information with different precision levels. For example, fighting ability can be influenced by prior experiences (Kar et al. 2016; Briffa and Lane 2017; Camerlink et al. 2017) and experienced individuals might be more capable of gleaning pertinent information during a fight. In this case, there may be a tendency for winners to have access to more precise information than losers.

Assessment strategies are implicitly treated as mutually exclusive within a contestant, but this might not be the case. An individual might adapt its strategy from fight to fight, but even within a fight both empirical and theoretical research has shown that exclusive use of a single strategy is not always the case (Prenter et al. 2006; Smallegange et al. 2007; Elias et al. 2008; Arnott and Elwood 2009). Contestants may use a strategy that is a mix of mutual and self-assessment. In this sense, we can imagine that contestants weigh their own RHP against the perceived RHP of their opponent. This ratio of information on self to opponent RHP would be 1:0 for pure SA and 1:1 for MA. Thus, a ratio of 2:1 would represent a scenario where a contestant uses opponent information, but

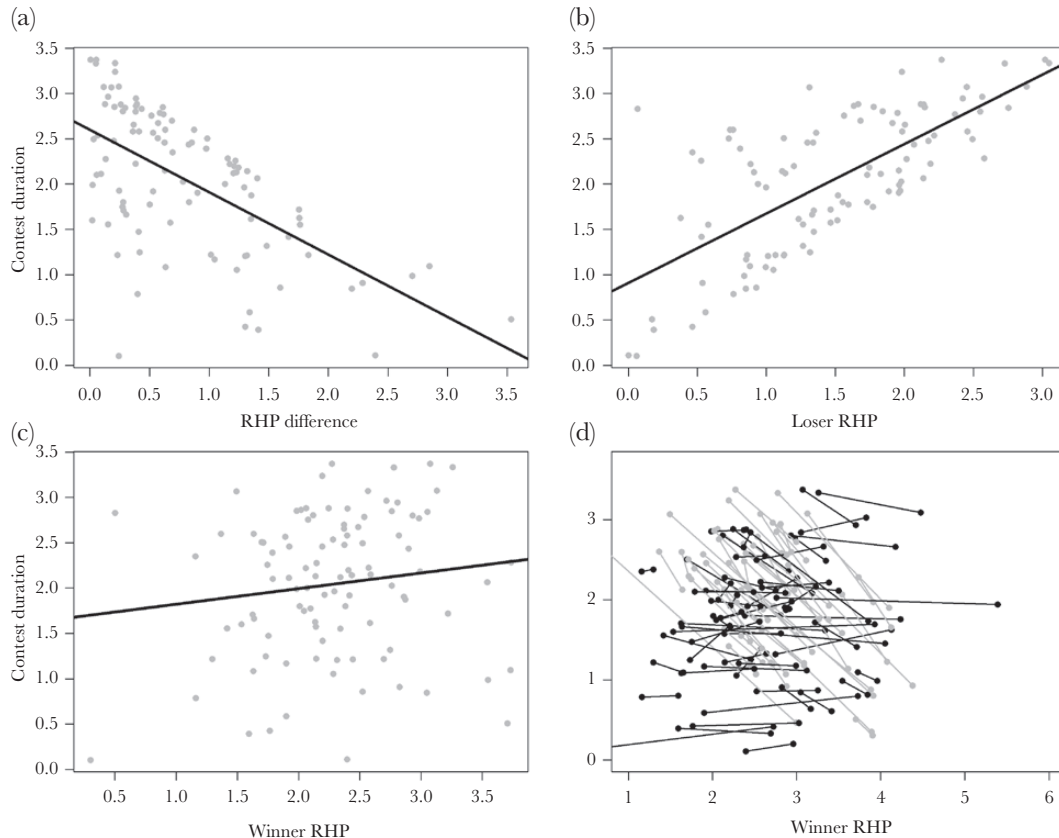


Figure 1

We used standard assessment strategy statistics with simulated data to illustrate limitations and problems with previous approaches. We randomly assigned RHP values to paired agents from a normal distribution such that $\text{mean} \pm \text{SD} = 0 \pm 1$ for 100 groups of three individuals and made contest duration correlate with either the RHP difference between dyadic opponents, or only the loser RHP. For this simulation we considered a population in which 50% of the individuals perform mutual assessment and 50% self-assessment. Researchers used to evidence mutual assessment via a negative correlation between the difference in RHP with contest duration (a) — an approach shown to be inappropriate (Taylor and Elwood 2003). Instead, loser and winner RHP should be examined separately (b, c). Using the simulated data, we see that loser, but not winner, RHP is correlated with contest duration. We show that this test is also insufficient for these data, as it assumed that assessment strategies are disjoint and fixed. Instead, we employ a repeated-testing approach (d), which reveals that our simulated population is actually 50% self and 50% mutual assessing. We assigned individuals as mutual or self-assessing if the slope was more (gray lines) or less than (black lines) the upper limit of the 95% confidence interval off all negative slopes. This approach accurately assigned around 80% of individual assessment strategies with a 0.25 error term (see Supplementary Model Code for a detailed description). In summary, a incorrectly supports mutual assessment, b and c incorrectly support self-assessment, and d reveals the true pattern of mixed assessment strategies in the population.

not to the same extent that it relies on its own RHP (Rillich et al. 2007). This has been termed partial MA (Prenter et al. 2006), such that one can envisage variation in assessment strategies as a continuum between the extremes of pure SA and assessment of the opponent's RHP.

CAM is a type of SA model because (as in the energetic war of attrition [EWOA]) individuals base their giving up decision on accumulated cost thresholds. In CAM, however, these costs come from two sources. First (as in the EWOA), there are the energetic costs of performing agonistic behavior. Added to these are the costs of injuries that the opponent has inflicted (Payne 1998). This, however, is an incomplete picture of the contest, because the rate at which an individual can inflict injuries on its rival (or the severity of the injuries that it inflicts) should be related to its RHP. In this sense, although displays that transmit information on RHP are absent from the CAM, it is conceivable that a recipient of injuries could make inferences about its rival's RHP, based on those injuries that it has received. Regardless of whether the weaker opponent glean information on rival RHP in this way, the relation between RHP and the ability to inflict injuries alone is enough to ensure that (at the very least) contest duration should vary as a function of both winner and loser RHP. Specifically, contests should correlate positively with loser RHP and negatively with winner RHP. This, of course, renders fights described by the CAM indistinguishable from fights settled by MA, under the Taylor-Elwood approach (Taylor and Elwood 2003; Briffa and Elwood 2009).

EXPANDING THE THEORETICAL FRAMEWORK

Individuals can potentially assess three components relevant to animal contests: themselves, their opponent, and the contested resource. While this has been acknowledged by researchers (Arnott and Elwood 2009; Keil and Watson 2010; Elwood and Arnott 2012; Fawcett and Mowles 2013), no framework has been developed that simultaneously incorporates these three components, and these considerations have yet to be adopted by empiricists. The three variables might vary in the attainability and accuracy of information, but also the relative weight placed on each information source. This can be thought of as a ternary diagram (Figure 2). In this sense, animal contest assessment has three extremes, within which all scenarios of animal contests lie: only use information about themselves, only use information about the opponent, or only use information about resource.

Among the three pure assessment strategies depicted in the ternary diagram, only opponent assessment has never been formally modeled. Therefore, it is not clear under what scenarios we might see the evolution of opponent assessment. A starting point might be organisms that cannot reliably estimate their own RHP. An understanding of one's own relative RHP (i.e., how one's RHP compares to the population) can be either learned or intrinsic. Animals might not be able to learn their relative RHP if they rarely compete with conspecifics, are young with little contest experience, or if the population demographics have changed. Intrinsic physiological or morphological cues might be unreliable if age or developmental stage is a poor correlate of RHP. Further, opponent information may not need to be compared to a reference. Such a scenario may be hard to imagine, but (intuitively) so is pure SA, where animals gain no information from opponents. Indeed, pure opponent assessment has been noted (Rillich et al. 2007; Prenter et al. 2008; Arnott and Elwood 2010; Reddon 2011). Thus, we posit that

opponent-dominant and self-dominant assessment strategies should be considered, with MA actually representing a midpoint along this spectrum.

The third pure strategy is that contestants might only use information about the resource (food, shelter, or mates), and not themselves or their opponent. This might occur if, for example, resources are highly valuable, such that invading opponents fight maximally regardless of self or opponent RHP. Such a scenario would appear similar to SA under the standard approach of correlating winner RHP with contest duration. This is because SA, under the current approach, is diagnosed by a lack of negative correlation between winner RHP and contest duration. Further, resource value interacts with RHP, which can be delineated into either physical or motivational RHP (Chapin and Hill-Lindsay 2016). Indeed, it may be the case that inconclusive studies under the Taylor-Elwood approach might be explained by resource assessment (e.g., Rudin and Briffa 2011).

Motivation matters when there is an asymmetry in resource value, resulting in one contestant more motivated than the other to win a contest (Humphries et al. 2006; Bergman et al. 2010; Santos and Peixoto 2017). This is commonly expressed as a residency effect (Olsson and Shine 2000; Kemp and Wiklund 2004; Kasumovic et al. 2011; Chapin and Hill-Lindsay 2016; Page and Coates 2017), but could occur for a number of reasons, like differences in satiation, resource distribution, information quality, or age (Sneddon et al. 2003; Brown et al. 2006; Humphries et al. 2006; Arnott and Elwood 2008; Petersen et al. 2010). Therefore, a more complete evaluation about the rules that animals use to settle contests should incorporate both information on traits linked to RHP

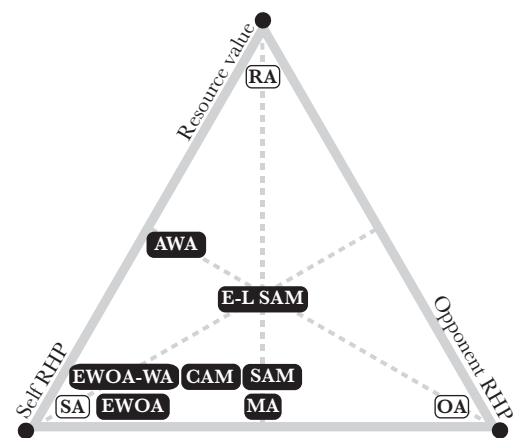


Figure 2

Ternary diagram of the animal contest framework. Corners indicate the three information sources used in contests. The weights placed on each information source determines the assessment strategy and game theory model, such that individuals using only one information source are using a “pure” strategy and any combination is a mixed strategy. SA is pure self-assessment; OA is pure opponent assessment; RA is pure resource assessment. Mutual assessment (MA) and the models that make this assumption (e.g., Sequential Assessment Model, SAM; Enquist et al. 1990) fall on the opponent-self information continuum. Asymmetric War of Attrition (AWA; Parker and Rubenstein 1981) falls on the self-resource spectrum, and Enquist-Leimar Sequential Assessment Model (E-L SAM) employs all three information sources. Dotted lines indicate the midpoint between pure strategies. Models that do not consider all three information sources still have utility in describing assessment strategies, but studies of animal contest should consider all information sources to enable hypothesis testing of the importance of all three information sources.

and measures (or controls) for the value that each individual gives to the disputed resource.

The influence of extrinsic factors beyond opponent RHP and resource value have rarely been examined in the context of animal contests. We have identified six factors that, while not a direct component of contests, can influence both assessment strategies and contest outcomes. These include the distribution of resources, population demography, individual experience, the quality of information transfer, ontogeny, and the expectation of future reproductive success. Each of these areas offers new avenues of research to incorporate additional factors into contest dynamics.

The distribution of resources can change optimal assessment strategies. Empirical data on how resource distribution affects contests remain scarce, but there is evidence, for example, that individuals in triadic interactions might escalate less because benefits might be reduced when resources are shared (Moore and Greeff 2003). Future research that experimentally alters resource distributions would be a valuable contribution.

Demographic patterns of RHP and how they change over time can shape optimal assessment strategies. Populations with generally stable RHP distributions might evolve innate thresholds for contests, while populations with frequent stochastic variation in demography might rely more on assessing opponents.

Individuals often show within-population variation that is consistent across time and contexts. Polymorphic crickets, for example, show distinct assessment strategies between morphotypes (Zeng et al. 2016). Nowbahari et al. (1999) found that large desert ants (*Cataglyphis niger*) use MA, while small individuals use SA. Such stable differences in behavior are usually referred to as animal personalities. For example, boldness describes an individual's propensity to take risks and aggressiveness is the propensity to initiate and escalate a contest (Briffa et al. 2015). Recent studies indicate that the two axes of variation might form a behavioral syndrome, where bolder individuals are also more aggressive. In sea anemones, for example, individuals that show short startle responses also land more blows on the opponent during a fight (Rudin and Briffa 2012). Thus, injuries may be more important in contests with bold individuals. Since injuries provide an individual with information about its own state, SA might be more important in fights between bolder individuals even if MA is possible in fights that occur between less bold members of the same population.

Although there has been some debate about the relative cognitive demands of SA versus MA (see Reichert and Quinn 2017), it is often assumed that MA is the more cognitively demanding task of the two (Elwood and Arnott 2012). Thus, if consistent intraspecific variation in cognition is present (see Griffin et al. 2015 for a discussion), this may influence the use of assessment rules during combat. For instance, individuals of greater cognitive ability might be better able to incorporate two sources of information (e.g., self and opponent RHP) into their decision making whereas individuals that fall lower on the cognitive range might rely to a greater extent on single sources of information and thus be more likely to use self-assessment or opponent assessment (or indeed, resource only assessment). Variation in decision rules could have cognitive causes, but this remains to be shown empirically.

Past contest experiences can influence not only the RHP of opponents, but also the assessment strategy that they employ. Previous victories and losses can have short-term impacts on motivation, and thus, RHP (Dugatkin 1997; Hsu and Wolf 1999). These effects usually last minutes to days, depending on the taxa. Beyond winner and loser effects, past experiences can inform individuals of

the RHP population distribution and how their own RHP relates to the population. More experienced individuals may be more efficient in evaluating their rivals and consequently may rely more on a mutual than self-assessment strategy (e.g., Elias et al. 2008). Past experiences are not requisite for assessment, however, as contest decisions can be determined innately.

Information accuracy is critical for animal contests. Information with high variance, erroneous signals, or cheating displays can all shape contest dynamics by reducing signal reliability. A lower reliability may favor strategies mainly based on SA, since individuals that rely on mutual information transfer may pay higher costs via wrong estimations (e.g., dishonest signaling; Chapin and Reed-Guy 2017). Residents generally have better information about the resource they are defending than putative usurpers (Rosenberg and Enquist 1991; Peixoto and Benson 2012; Chapin and Hill-Lindsay 2016). Further, winners in particular, or more experienced individuals in general, might have better information than losers or naive opponents (McCallum et al. 2016; Briffa and Lane 2017). As mentioned earlier, hermit crabs usurping shells have good information about the contested resource but poor information about the opponent's RHP, while the opposite is true for defenders (Briffa and Elwood 2004).

Lastly, ontogenetic effects can shape assessment strategies. There is some empirical support for this, where opponent size not only predicts contest outcomes, but also assessment strategy (Constant et al. 2011). Another possibility involves the residual reproductive value (Williams 1966). Older individuals may have lower reproductive expectation than younger ones (i.e., lower residual reproductive value). In this scenario, older opponents may pay lower costs in terms of future reproductive success lost if they increase their investment in a fight (Kemp 2006). Therefore, in some species, individuals may become more aggressive as they age (e.g., Kemp 2002) and this may affect the assessment strategy adopted by them.

Many of the additional factors that potentially affect the assessment strategies mentioned here may be controlled for or included in analyses of animal contests. In particular, past experiences and age should be experimentally controlled or included in statistical analyses, while possible asymmetries in the amount of information about the resource value and the reliability of information transfer must be investigated before examining the relationship between RHP traits and contest duration. For instance, to control for winner/loser effects, trials could be planned far enough apart for winner and loser effects to diminish, while individuals with similar age and previous experiences may be paired in fighting trials. Most studies on the duration of winner and loser effects suggest that they do not last long (more than a week in some species, but less than an hour in others), but we expect this to vary with life history and social structure, including population demography and density (Hsu et al. 2006; Huang et al. 2010; Kasumovic et al. 2010). Second, winner/loser and ontogenetic/experience effects can be controlled by using the proportion of fights won or age and the number of prior fights as covariates in a multiple-individual model, such as a generalized linear model.

EMPIRICAL GUIDELINES

We recommend, where logistically possible, a repeated-trials approach to test assessment strategy models. This approach does not assume that assessment strategies are binary, that all individuals in a population share the same strategy, or that strategies are ontogenetically fixed within individuals. First a suitable proxy for RHP should be identified by investigating measurable differences

between contest winners and losers of the study species (e.g., body size, weight, weaponry, behavior; Rudin and Briffa 2011; Briffa 2014). Next focal animals are allocated to multiple opponents of a variety of RHP levels, but all of higher RHP than the focal, and in random order. The extent to which the focal animal uses a mutual or self-assessment strategy is indicated by the slope of a best-fit line of opponent RHP and contest durations (e.g., Figure 1). Since the focal animal will lose in most trials, a negative slope indicates that the RHP of the winner affects contest duration, evidencing opponent assessment. Slopes approaching zero indicate a lack of opponent assessment. By this method, assessment strategies can be compared among species, populations, and individuals within populations. The ability to examine among-individual variation in assessment strategy is an inherent feature of our proposed approach—a much discussed, but hitherto untested area of animal contest research.

Note that this approach does not account for mid-contest strategy switching. Researchers should test for assessment strategies across escalation-based stages of a contest. For example, Hsu et al. (2006) divided their analysis of assessment strategies into contest stages and found that killifish (*Kryptolebias marmoratus*) use MA during initial phases of a contest, but switch to SA after the contest has escalated. Adding repeated trials to this sort of experimental design would also reveal among-individual variation in assessment strategies.

Further, researchers can compare treatments of resources that vary in value, with both the focal or opponent as the resource holder. This will disentangle the role of variable resource value and information between opponents in dictating contest duration and escalation.

To evaluate the efficiency of our approach in identifying the assessment rule adopted by each individual, we built a simulation model (see Supplementary Material). We randomly chose RHP values from a normal distribution ($\bar{x} = 0$, standard deviation [SD] = 1) and randomly assigned RHP values to individuals grouped into 100 triads. We then normalized RHP values to remove negative numbers and assigned individuals with the smallest RHP value of each triad to act as the focal individual, and the other two, larger individuals, as opponent individuals. We randomly assigned focal individuals to use SA or MA across five model versions that varied in the proportion of mutual to self-assessors: 1:0, 0.3:0.7, 0.5:0.5, 0.7:0.3, and 0:1.

We calculated contest duration as the giving up time of the losing focal individual. For SA, we calculated this as the RHP of the focal individual plus an error term (because individuals generally have imperfect information), such that

$$\min \{D_1 \dots D_n\} + RHP_{focal} + error$$

Where D is the duration for n individuals in the simulation. We added the minimum value of fight duration to the giving up time of each individual performing SA to avoid negative values of fight duration. Organisms using MA, however, should shorten their giving up time in light of opponent size. As such, we calculated the giving up time for mutually assessing agents in our model as,

$$\max \{D_1 \dots D_n\} - |RHP_{focal} - RHP_{opponent}| + error$$

or the maximum duration across all trials minus the absolute value of opponent contest durations for a given trial, plus an error term. For both mutual and self-assessment, we randomly selected

the error term from a uniform distribution $U(a, b)$, such that, $a = -1 \times b$.

We used the simulated data to illustrate the limitations of current approaches and the efficacy of ours. If we correlate the difference in opponent RHP with contest duration (considering only the first fight for each focal loser), we find a negative correlation for all five model versions (Table 1; Figure 1a). As pointed out in Taylor and Elwood (2003), this approach could lead to the unreliable conclusion that most losers are using MA. If we use the alternative approach suggested by Taylor and Elwood (2003) and examine the relationship between RHP and contest duration for opponents separately, we find that loser RHP positively correlates with contest duration for all model versions (Table 2, Figure 1b). The winner, on the other hand, positively correlated with contest duration when 100% of the population was composed of self-assessors and did not correlate with contest duration for scenarios that varied from 70% to 30% of self-assessors (Table 2, Figure 1c). A negative correlation between contest duration and winner RHP occurred only when all individuals in the population adopted MA (Table 2). Therefore, the results using the Taylor and Elwood approach for scenarios that varied from 0.3 to 0.7 of the population adopting MA also led to misleading conclusions.

We advocate two approaches to disentangle populations with mixed contest assessment: to test if populations are composed of mixed strategies inferentially, and to estimate individual-level assessment strategies. The presence of mixed strategies can be tested inferentially under the assumption that slopes should be near zero if individuals use SA.

We can also estimate the assessment strategies of individuals with our repeated trials approach. Without error, the slopes of opponents using SA and contest duration should be zero and MA slopes should be negative. However, in natural systems, experimental error will alter slopes, and an alternative cutoff may be preferred. We propose a threshold calculated as the mean plus 95% CI for negative slopes (the expectation under MA). We assigned an individual's strategy as MA if the slope was smaller than the upper limit of the confidence interval and SA if larger. We ran 20 simulations with this threshold to illustrate assignment accuracy for each model

Table 2

Linear regressions of simulated data predicting contest duration illustrating different approaches used to investigate animal contest assessment strategies ($n = 100$)

Predictor	MA:SA	F	r ²	P	m
RHP difference	1:0	2718	0.96	<0.001	-
	0.7:0.3	235.1	0.7	<0.001	-
	1:1	88.77	0.47	<0.001	-
	0.3:0.7	12.73	0.11	<0.001	-
	0:1	60.46	0.38	<0.001	-
loser RHP	1:0	40.56	0.29	<0.001	+
	0.7:0.3	40.94	0.29	<0.001	+
	1:1	95.83	0.49	<0.001	+
	0.3:0.7	41.09	0.30	<0.001	+
	0:1	3026	0.97	<0.001	+
winner RHP	1:0	21.42	0.17	<0.001	-
	0.7:0.3	1.12	0.001	0.30	+
	1:1	0.20	<0.001	0.66	+
	0.3:0.7	5.39	0.04	0.02	+
	0:1	32.72	0.24	<0.001	+

MA, mutual assessment; SA, self-assessment; m is slope sign among different simulations.

version (Table 3, Figure 1d). The mean accuracy varied between 76% and 85% and between 70% and 87% when considering the efficiency for mutual or SA separately (with most model versions showing efficiency above 80%). It is important to note that the accuracy of both approaches is a function of experimental error (included in our model as an error term). Removing the error term results in estimate accuracy reaching 100%.

A focal losing twice is the minimum requirement for testing individual-level assessment strategies. We explored this type of data here in our simulation as a greater number of fights may not be feasible in many real study systems (due to accumulated injuries for instance). Where feasible, adding additional opponents could reveal important within-individual variation across contests and contexts. If each focal fought multiple times, one could employ random regression models, where a significant random slope effect would indicate the presence of different responses to opponent RHP across focal individuals. Once such a focal \times opponent interaction (a type of behavioral reaction norm, analogous to the more familiar “individual \times environment” interaction; see Dingemans et al. 2010) was revealed, those individual slopes could be examined in more detail. In our sample simulation, this would only describe the random error we added, but in natural populations, variation in assessment is likely important. For example, repeated contests where resource value is manipulated, or information about resource value is constrained, could reveal how resource value influences contests. Further, these analyses can be conducted on observational data of random and naturally occurring contests; data need not be from staged laboratory contests. However, naturally occurring contests may include nonrandom pairings. For example, contests can be more size-matched than expected (Fawcett and Mowles 2013) or include opponents with differences in motivations to compete (e.g., Bergman et al. 2010; Santos and Peixoto 2017). This can make it difficult for researchers to ascertain which opponent is of higher RHP, or which won or lost the contest.

The new approach facilitates meta-analyses, which could identify generalities in within-population contest assessment strategies, including the putative existence of pure opponent, self, or other strategies that involve a mixture of information sources. In particular, researchers should report focal individual-level measures of slope, variance, the coefficient of determination, and also the same measures pooled for all individuals.

We propose a new framework for animal contests focusing on three key sources: the individual, its opponent, and the contested resource. The relative weights that individuals put on these three information sources identifies the assessment strategy employed by individuals, and practically all models of assessment strategy fit within this parameter space. This highlights that, despite their historical importance, the existing models of contest resolution rules

Table 3
Mean \pm SD percent correct assignment of assessment to 100 focal individuals using the individual-level slopes approach

MA:SA	Overall	MA	SA
0.7:0.3	75.5 \pm 7.9	70.4 \pm 12.6	87.5 \pm 21.5
0.5:0.5	79.2 \pm 12.8	81.8 \pm 9.6	76.5 \pm 31.5
0.3:0.7	85.2 \pm 11.0	87.5 \pm 6.2	84.1 \pm 16.5

Mean percentages were based on 20 simulations. MA:SA indicates the proportion of individuals adopting mutual and self-assessment in the population.

do not cover all assessment possibilities that may occur in fights. Assessment strategies might not be fixed within or between populations, individuals, or contests, and our understanding of fighting would benefit from experiments that can identify a variation in assessment strategies where this is present. Further, contest dynamics are affected by past experiences, future prospects, population demography, resource distributions, information quality, and ontogeny. MA, in particular, may be viewed as a mid-range along a continuum of self and opponent assessment. More importantly, the possibility that the relative importance of self and opponent assessment may vary within and between individuals should be investigated. Individual-level repeated testing can be a useful tool in this regard, although, for a complete understanding of how rivals behave during contests, additional information is needed. Moving forward, empirical work on animal contests should focus on individual-level variation within populations and the mechanisms that promote not just a binary choice between strategies for contest resolution, but more dynamic strategies that vary across populations, time, and contexts. Similarly, theoretical work is needed to understand the evolutionary dynamics of assessment strategy variation in populations and individuals.

SUPPLEMENTARY MATERIAL

Supplementary data are available at *Behavioral Ecology* online.

FUNDING

P.E.C.P. was supported by grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq produtividade em pesquisa: 305561/2014-6).

Editor-in-Chief: Leigh Simmons

REFERENCES

- Arnott G, Elwood RW. 2007. Fighting for shells: how private information about resource value changes hermit crab pre-fight displays and escalated fight behaviour. *Proc Biol Sci.* 274:3011–3017.
- Arnott G, Elwood RW. 2008. Information gathering and decision making about resource value in animal contests. *Anim Behav.* 94:529–542.
- Arnott G, Elwood RW. 2009. Assessment of fighting ability in animal contests. *Anim Behav.* 77:991–1004.
- Arnott G, Elwood RW. 2010. Startle durations reveal visual assessment abilities during contests between convict cichlids. *Behav Processes.* 84:750–756.
- Bergman M, Olofsson M, Wiklund C. 2010. Contest outcome in a territorial butterfly: the role of motivation. *Proc Biol Sci.* 277:3027–3033.
- Bishop DT, Cannings C. 1978. A generalized war of attrition. *J Theor Biol.* 70:85–124.
- Brandt Y, Swallow JG. 2009. Do the elongated eye stalks of diopsid flies facilitate rival assessment? *Behav Ecol Sociobiol.* 63:1243–1246.
- Briffa M. 2008. Decisions during fights in the house crickets, *Acheta domesticus*: mutual or self assessment of energy, weapons, and size? *Anim Behav.* 75:1053–1062.
- Briffa M. 2014. What determines the duration of war? Insights from assessment strategies in animal contests. *PLoS One.* 9:e108491.
- Briffa M. 2015. Agonistic signals: integrating analysis of functions and mechanisms. In: Irschick DJ, Briffa M, Podos J, editors. *Animal signaling and function, an integrative approach.* Hoboken (NY): Wiley. doi:10.1002/9781118966624.ch6
- Briffa M, Elwood RW. 2004. Use of energy reserves in fighting hermit crabs. *Proc Biol Sci.* 271:373–379.
- Briffa M, Elwood RW. 2009. Difficulties remain in distinguishing between mutual and self-assessment in animal contests. *Anim Behav.* 77:759–762.

- Palaoro AV, Briffa M. 2017. Weaponry and defenses in fighting animals: how allometry can alter predictions from contest theory. *Behav Ecol*. 28:328–336.
- Palaoro AV, Dalosto MM, Costa JR, Santos S. 2014. Freshwater decapod (*Aegla longirostris*) used a mixed assessment strategy to resolve contests. *Anim Behav*. 95:71–79.
- Parker GA. 1974. Assessment strategy and the evolution of fighting behaviour. *J Theor Biol*. 47:223–243.
- Parker GA, Rubenstein DI. 1981. Role assessment, reserve strategy, and acquisition of information in asymmetric animal conflicts. *Anim Behav*. 29:221–240.
- Payne RJH. 1998. Gradually escalating fights and displays: the cumulative assessment model. *Anim Behav*. 56:651–662.
- Payne RJH, Pagel M. 1996. Escalation and time costs in displays of endurance. *J Theor Biol*. 183:185–193.
- Payne RJH, Pagel M. 1997. Why do animals repeat displays? *Anim Behav*. 54:109–119.
- Peixoto PEC, Benson WW. 2012. Influence of previous residency and body mass in the territorial contests of the butterfly *Hermeuptychia fallax* (Lepidoptera: Satyridae). *J Ethology*. 30:61–68.
- Petersen A, Nielsen KT, Christensen CB, Toft S. 2010. The advantage of starving: success in cannibalistic encounters among wolf spiders. *Behav Ecol*. 21:1112–1117.
- Pinto NS, Palaoro AV, Peixoto PEC. 2019. All by myself? Meta-analysis of animal contests shows stronger support for self than for mutual assessment models. *Biol Rev*. 94:1430–1442.
- Pratt AE, McLain DK, Lathrop GR. 2003. The assessment game in sand fiddler crab contests for breeding burrows. *Anim Behav*. 65:945–955.
- Prenter J, Elwood RW, Taylor PW. 2006. Self assessment by males during energetically costly contests over precopula females in amphipods. *Anim Behav*. 72:861–868.
- Prenter J, Taylor P, Elwood RW. 2008. Large body size for winning and large swords for winning quickly in swordtail males, *Xiphophorus helleri*. *Anim Behav*. 75:1981–1987.
- Reddon AR, Voisin MR, Menon N, Marsh-Rollo SE, Wong MYL, Balshine S. 2011. Rules of engagement for resource contests in a social fish. *Anim Behav*. 82:93–99.
- Reichert MS, Quinn JL. 2017. Cognition in contests: mechanisms, ecology, and evolution. *Trends Ecol Evol*. 32:773–785.
- Rillich J, Schildberger K, Stevenson PA. 2007. Assessment strategy of fighting crickets revealed by manipulating information exchange. *Anim Behav*. 74:823–86.
- Rosenberg RH, Enquist M. 1991. Contest behaviour in Weidemeyer's admiral butterfly *Limenitis weidemeyerii* (Nymphalidae): the effect of size and residency. *Anim Behav*. 42:805–811.
- Rudin FS, Briffa M. 2011. The logical polyp: assessments and decisions during contests in the beadlet anemone *Actinia equina*. *Behav Ecol*. 22:1278–1285.
- Rudin FS, Briffa M. 2012. Is boldness a resource-holding potential trait? Fighting prowess and changes in startle response in the sea anemone, *Actinia equina*. *Proc Biol Sci*. 279:1904–1910.
- Santos TB, Peixoto PEC. 2017. Agonistic interactions in the dragonfly *Micrathyrus unguilata*: does male fighting investment come from an innate ability or an indomitable will? *Behav Ecol Sociobiol*. 71:104.
- Schnell AK, Smith CL, Hanlon RT, Harcourt R. 2015. Giant Australian cuttlefish use mutual assessment to resolve male-male contests. *Anim Behav*. 107:31–40.
- Sih A, Bell A, Johnson JC. 2004. Behavioral syndromes: an ecological and evolutionary overview. *Trends Ecol Evol*. 19:372–378.
- Smallegange IM, Sabelis MW, Meer J. 2007. Assessment games in shore crab fights. *J Experim Marine Biol Ecol*. 351:255–266.
- Sneddon LU, Huntingford FA, Taylor AC, Clare AS. 2003. Female sex pheromone-mediated effects on behavior and consequences of male competition in the shore crab (*Carcinus maenas*). *J Chem Ecol*. 29:55–70.
- Stuart-Fox D. 2006. Testing game theory models: fighting ability and decision rules in chameleon contests. *Proc Royal Soc London B*. 273:1555–1561.
- Taylor PW, Elwood RW. 2003. The mismeasure of animal contests. *Anim Behav*. 65:1195–1202.
- Tedore C, Johnsen S. 2015. Visual mutual assessment of size in male *Lyssomanes viridis* jumping spider contests. *Behav Ecol*. 26:510–518.
- Tibbetts EA, Mettler A, Levy S. 2010. Mutual assessment via visual status signals in *Polistes dominulus* wasps. *Biol Lett*. 6:10–13.
- Tsai Y-J, Barrows EM, Weiss MR. 2014. Why do larger and older males win contests in the parasitoid wasp *Nasonia vitripennis*? *Anim Behav*. 91:151–159.
- Williams GC. 1966. Natural selection, the cost of reproduction and a refinement of Lack's principle. *Am Nat*. 100:687–690.
- Wofford SJ, Earley RL, Moore PA. 2015. Evidence for assessment disappears in mid-sex contests of the crayfish, *Orconectes virilis*. *Behav*. 152:995–1018.
- Yasuda CI, Koga T. 2016. Importance of weapon size in all stages of male-male contests in the hermit crab *Pagurus minutus*. *Behav Ecol Sociobiol*. 70:2175–2183.
- Yasuda CI, Takeshita F, Wada S. 2012. Assessment strategy in male-male contests of the hermit crab *Pagurus middendorffii*. *Anim Behav*. 84:385–390.
- Zeng Y, Zhi D-H, Kang W-N. 2016. Variation in fighting strategies in male wing-dimorphic crickets (Gryllidae). *Behav Ecol Sociobiol*. 70:429–435.