Assessment of fighting ability in animal contests

Gareth Arnott*, Robert W. Elwood

School of Biological Sciences, Queen's University Belfast

Selection should favour accurate information gathering regarding the likely costs and benefits of continued conflict. Here we consider how variation in the abilities of contestants to assess resource-holding potential (RHP) influences fights. This has been examined in various game theory models. However, discriminating between assessment strategies has proven difficult and has resulted in confusion. To add clarity, we group existing models into three main types that differ in the information about RHP that contestants are presumed to gather: (1) pure self-assessment, (2) cumulative assessment and (3) mutual assessment. Within this framework we outline methods advocated to discriminate successfully between the three main assessment models. We discuss support for each model, before highlighting a number of conflicting and inconclusive studies, leading us to consider alternative approaches to investigate assessment. Furthermore, we examine support for newly emerging concepts such as 'varying degrees of assessment', 'switching assessment' strategies and the possibility of contestants adopting different assessment strategies within a fight involving distinctive roles. We suggest future studies will benefit by judicious use of a battery of techniques to determine how animals settle contests. Finally, we highlight difficulties with current game theory models, and raise concerns regarding the use of certain behavioural criteria to accept or reject a model, particularly since this may conflict with evidence for a given assessment strategy. Furthermore, the failure of existing models to account for newly emerging concepts points to limitations of their use and leads us to challenge game theoreticians to develop upon them.

© 2009 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

Fighting for limited resources may be costly, in terms of energy use, time, risk of injury and increased risk of predation or death (e.g. Glass & Huntingford 1988; Kelly & Godin 2001; Briffa & Elwood 2004) and selection should favour accurate information gathering, regarding the likely costs and benefits of continued conflict, to enable appropriate tactical decisions (Parker 1974; Maynard Smith & Parker 1976; Parker & Rubenstein 1981). How the ability to gather information about resources influences fight behaviour has recently been reviewed (Arnott & Elwood 2008). Here we consider how variation in overall fighting ability of each contestant, generally termed resource-holding potential (RHP), and the assessment abilities of the contestants, influence how each fights (Parker 1974).

Victory tends to go to the larger or heavier contestant so body size is often used as a proxy for RHP (see Table 1), since size is generally correlated with strength and the ability to inflict injury (Archer 1988). However, sometimes the smaller contestant wins, particularly when the difference in body size is small (Faber & Baylis 1993; Hughes 1996). Other correlates of fighting ability are also likely to differ between contestants, and include experience, position, development of weaponry, physiological state and sex, all of which may affect the chances of winning an encounter (see Table 1). Indeed, when multiple traits influence fighting ability, body size may not be the best indicator of an individual's fighting ability. For example, in cape dwarf chameleons, *Bradypodion pumilum*, contest outcome is influenced by the height of the ornamental casque, the relative size of the pink patch in the centre of the flank and previous experience (Stuart-Fox & Whiting 2005; Stuart-Fox et al. 2006) and Stuart-Fox (2006) calculated a 'multivariate' estimate of relative fighting ability (RHP). This 'multivariate' measure of RHP is potentially a very useful advance because it is more likely to predict outcome.

How RHP influences fights has been examined in various game theory models that consider contests settled by persistence. Here we group them into three main types that differ in the information about RHP that contestants are presumed to gather.

(1) Pure self-assessment: these are models in which each contestant only has information about its own abilities or state and fails to gather information about its opponent, and the actions of the opponent do not inflict costs, although both opponents incur a cost from their own actions. In this scenario, rivals persist purely in accord with their own RHP such that weaker rivals tend to reach their limits and give up first. This strategy, termed 'self-assessment',

* Correspondence: G. Arnott, School of Biological Sciences, Queen's University Belfast, Medical Biology Centre, 97 Lisburn Road, Belfast BT9 7BL, U.K.
   E-mail address: garnott01@qub.ac.uk (G. Arnott).

Table 1
Examples of correlates of fighting ability

Asymmetry	Effect of correlate on contests	Source
Body size		
Bowl and doily spider	Larger animals typically win contests	Austad 1983
Frontinella pyramitela		
Stomatopod crustacean		Adams & Caldwell 1990
Gonodactylus bredini		
Sand goby		Lindstrom 1992
Pomatoschistus minutus		
Convict cichlid		Draud & Lynch 2002
Archocentrus nigrofasciatum		
Jumping spider		Wells 1988
Euophrys parvula		
Lizard		Stamps & Krishnan 1994
Anolis aeneus		
Net-spinning caddis larva		Englund & Olsson 1990
Arctopsyche ladogensis		
Body mass		
Desert spider	Heavier animals typically win contests	Riechert 1978
Agelenopsis aperta		
Pig		Rushen 1987
Sus scrofa		
Pumpkinseed sunfish		Dugatkin & Ohlsen 1990
Lepomis gibbosus		
Tree lizard		Zucker & Murray 1996
Urosaurus ornatus		
Red deer		Clutton-Brock et al. 1979
Cervus elaphus		
Cichlid fish		Enquist et al. 1990
Nannacara anomala		
Experience		
Australian dragon lizard	Previous winners tend to win fights	Stuart-Fox & Johnston 2005
Ctenophorus decresii, C. vadnappa		
Tree lizard	Previous winners tend to win fights, previous losers tend to lose fights	Zucker & Murray 1996
Urosaurus ornatus		
Mangrove rivulus fish	As above	Hsu & Wolf 1999
Rivulus marmoratus		
Position		
Amphipods	Positional advantage to owners of precopula females in fights	Dick & Elwood 1990
Gammarus pulex		
Red-spotted newt	As above	Verrell 1986
Notophthalmus viridescens		
Development of weaponry		
Red deer	Number of antler points is weakly correlated with fighting success	Clutton-Brock et al. 1979
Cervus elaphus		
Shore crab	Crabs with longer chelae more likely to win fights	Sneddon et al. 1997
Carcinus maenas		
Wellington tree weta	Males with larger weapons (mandibles) more likely to win contests	Kelly 2006
Hemideina crassidens		
Prawns	Individuals with larger chelipeds more likely to win fights	Barki et al. 1997
Macrobrachium rosenbergii		
Dung beetle	Individuals with longer horns more likely to win fights	Pomfret & Knell 2005
Euoniticellus intermedius		
Physiological state		
Fruit fly	Males raised at higher temperature won more contests	Zamudio et al. 1995
Drosophila melanogaster		
African elephant	Musth males dominate nonmusth males	Poole 1989
Loxodonta Africana		
Damselfly	Fat reserves important: 'fatter wins rule'	Marden & Rollins 1994
Calopteryx maculata		
Speckled wood butterfly	Body temperature important: warmer individuals more likely to win contests	Stutt & Wilmer 1998
Pararge aegeria		
Butterfly	Age important: winners older than losers	Kemp 2000
Hypolimnas bolina		
Sex		
Great tit	Males defeat females in competitive interactions	Wilson 1992
Parus major		

and previously referred to as own RHP-dependent persistence or own size assessment (Taylor & Elwood 2003), is a feature of the 'war of attrition without assessment' (WOA-WA; Mesterton-Gibbons et al. 1996), and 'energetic war of attrition' (E-WOA; Payne & Pagel 1996, 1997) models, which make qualitatively similar predictions. Because of confusion with another form of self-assessment, the cumulative assessment model or CAM detailed below, we refer to these models as 'pure self-assessment'.

(2) Cumulative assessment: this category is encompassed by the 'cumulative assessment model' (CAM; Payne 1998), which is a game of both endurance and tolerance to damage inflicted by an opponent (Payne 1998). As above, this model is one of self-assessment with contestants terminating the contest when accrued costs exceed an absolute individual threshold, and no direct information is gathered about the opponent (Payne 1998). However, unlike the other self-assessment models, in which costs accrue only as a result of each rival's own behaviour, in the CAM costs also accrue from the opponent's actions, and superior opponents are better at inflicting costs. In other words, in the CAM the decision to withdraw is influenced by both an individual's own RHP (poor-quality individuals can bear fewer costs) and the opponent's RHP (higher quality individuals can inflict costs at a higher rate; Payne 1998).

(3) Mutual assessment: these are models in which both contestants assess their opponent's RHP relative to their own, with the selective advantage that the one with the lower RHP can quickly terminate the contest and thus reduce time, energy and risk of injury from engaging in a contest that it would inevitably lose. This strategy, termed 'mutual assessment' is central to the 'sequential assessment model' (SAM; Enquist & Leimar 1983; Enquist et al. 1990) and the 'asymmetric war of attrition' (AWOA) model of animal conflict (Parker & Rubenstein 1981; Hammerstein & Parker 1982). In the SAM, activities are performed in a series of phases that reveal information about fighting ability. Each phase can be thought of as a statistical sampling process (containing a degree of random error) and thus more 'sampling' provides a better estimate of contestant asymmetry (Enquist et al. 1990). Fights typically begin with low-cost/low-intensity elements that are relatively unreliable and, if the asymmetry is small, more 'sampling' will be required and hence the contest will proceed to the use of higher cost/higher intensity elements that better indicate RHP (Enquist et al. 1990).

Pure self-assessment and CAM are clearly inferior to mutual assessment as the contestant will always incur costs up to a threshold. However, basing decisions on individual thresholds (self-assessment) to determine contest duration may be an economical way to determine the degree of escalation, and, ultimately, contest winner, while avoiding the costs associated with rival assessment (Taylor & Elwood 2003; Elias et al. 2008). The time and energy required to process an opponent's cues and signals could be substantial, and detract from a contestant's own performance in a fight. The potential costs of assessment would be even more extreme if cues and/or signals are unreliable indicators of actual fighting ability (Elias et al. 2008). Prefight signals may at times exaggerate abilities (Hughes 2000; Elwood et al. 2006) and therefore reduce reliability of signal cues. Furthermore, the resource that is being fought over may be so valuable, in terms of future expected fitness, that 'no-assessment strategies' may be favoured (Moore et al. 2008).

The three broad hypotheses of assessment do not differ in predicting contest winners but in when and how the decision to give up is made (Prenter et al. 2006). Recently, Taylor & Elwood (2003) highlighted the fact that analyses used in previous studies failed to distinguish between them, and may have overestimated the extent to which animals engage in mutual assessment.

In cases of mutual assessment, it is predicted that contestants perceive large differences in RHP (e.g. size) more readily than small differences, producing a negative relation between RHP difference and contest duration (or other measures of contest cost). This negative relation has been one of the most frequently tested predictions in studies of animal conflict and is commonly cited as diagnostic of mutual assessment (e.g. Austad 1983; Enquist & Leimar 1983; Englund & Olsson 1990; Enquist et al. 1990; Leimar

et al. 1991; Rosenberg & Enquist 1991; Faber & Baylis 1993; Marden & Rollins 1994; Smith et al. 1994; Dale & Slagsvold 1995; Morris et al. 1995; Jennions & Backwell 1996; Mesterton-Gibbons et al. 1996; Hack 1997a; Renison et al. 2002; Pratt et al. 2003). Thus, most studies of how contest dynamics, duration and outcome are influenced by RHP have focused exclusively on a composite measure of RHP difference (e.g. larger – smaller, Stokkebo & Hardy 2000; smaller/larger, Pavey & Fielder 1996; ln(larger) – ln(smaller), Leimar et al. 1991; Faber & Baylis 1993; absolute size difference/smaller size, Dowds & Elwood 1985; Moya-Larano & Wise 2000; absolute size difference/larger size, Wells 1988; Dugatkin & Biederman 1991; absolute size difference/average size, Hack 1997a). The general conclusion is that because fight duration or cost is negatively related to the composite measure of RHP difference, the contestants must have gathered information about the opponents' RHP relative to their own (Taylor & Elwood 2003). However, there are significant problems with this approach, the details of which are outlined below.

A few studies have examined the individual contribution of each contestant's size to the contest duration and some of these have cast doubt on the idea of mutual assessment. For example, in *Metellina mengei* orb-web spiders, both intensity and duration of contests increased with the loser's size, whereas the winner's size appeared to be unimportant (Bridge et al. 2000). Also, maximum escalation and duration of contests between male *Plexippus paykulli* jumping spiders increased with absolute size of the smaller rival (usually the loser) in a pair but was not influenced by size of the larger rival (Taylor et al. 2001). Additionally, contests involving size-matched individuals are not expected to vary with absolute contestant size if mutual assessment occurs. However, Taylor et al. (2001) found a positive relation between size and both maximum escalation and duration. Similarly, in *Argyrodes antipodiana* spiders, escalation tendency of size-matched rivals is positively associated with body size (Whitehouse 1997) and in size-matched *Gryllus integer* crickets and *Liocarcinus depuratus* and *Uca annulipes* crabs contest duration is positively related to average size (Dixon & Cade 1986; Glass & Huntingford 1988; Jennions & Backwell 1996). Thus, the trends observed in these studies are consistent with self-assessment and are inconsistent with resolution through decisions of retreat based on estimated size (RHP) differences (mutual assessment; Taylor & Elwood 2003).

To examine the relations between contest duration (or other costs) and measures of individual contestant RHP, as well as composite measures of RHP difference, Taylor & Elwood (2003) simulated a population of individuals with varying RHPs (sizes). In the first simulation, each individual was set to persist for a duration that was exactly equal to its own RHP (size), and thus contest duration was entirely explained by the RHP of the loser, that is, by pure self-assessment. Although biologically unrealistic, this example provides the clearest illustration of how pure self-assessment can produce misleading 'effects' that have been inadvertently interpreted as showing mutual assessment.

Simple regression showed a strong positive relation (in this case perfect) between smaller (loser) rival size and contest duration (Fig. 1a), and the larger rival's size was also positively, but more weakly, related to duration (Fig. 1b). However, this simulation also showed that contest duration was negatively related to size (RHP) asymmetry (Fig. 1c). That is, a simulation of contests, the duration and outcome of which were settled purely by the size of the smaller contestant, nevertheless produced a negative relation between duration and any of the composite measures of RHP difference. Thus the sole use of a composite measure, and its relation to duration, to support the idea of mutual assessment was shown to be spurious because the same relation was obtained by pure self-assessment.
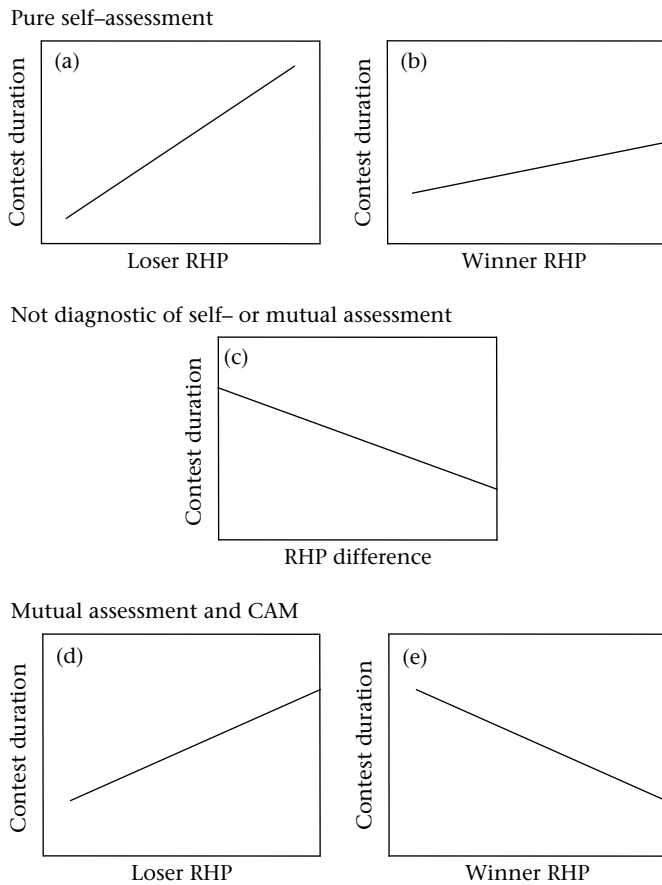
Pure self–assessment

Not diagnostic of self– or mutual assessment

Mutual assessment and CAM

Figure 1. Relations between resource-holding potential (RHP) and contest duration for alternative assessment hypotheses. (a, b) Pure self-assessment, (c) Not diagnostic of self- or mutual assessment, (d, e) Mutual assessment and cumulative assessment model. (a, d) Loser RHP, (b, e) Winner RHP. Adapted from Taylor & Elwood 2003.

This spurious relation is explained as follows: any individual may be matched randomly with a rival of similar strength, but only the weakest members of the population will encounter opponents that are much stronger (and vice versa). That is, the tendency for contests between rivals of very different RHP to be resolved quickly may be simply because, as RHP difference increases, the pool of weaker rivals potentially involved in a pairing becomes ever more restricted to the least persistent members of the population (Taylor & Elwood 2003). In other words, the apparent negative relation between contest duration and RHP (size) asymmetry in Taylor & Elwood's simulation is actually due to the (unavoidable) negative relation between RHP asymmetry and loser (smaller rival) RHP.

Taylor & Elwood (2003) also carried out a more biologically relevant simulation in which error (variation) was introduced and the relation between the persistence of an individual and its RHP was not exact. All trends were broadly similar to the case without variance. These relations can also be investigated by multiple regression, using larger and smaller rival size as covariates. Such an approach counters the incidental effects of larger rival size that arise through association with smaller rival size. In this case, Taylor & Elwood (2003) found the effects of larger rival size were reduced to nonsignificance. Similar results may be obtained in a related multiple regression using smaller rival size and size difference as covariates.

Taylor & Elwood (2003) also used a simulated population in which contest duration depended on the size difference between rivals (mutual assessment), and found that contest duration increased with smaller rival size and decreased to a similar extent (i.e. approximately opposite slopes) with larger rival size both in simple regression and when considered as covariates in multiple regression (Fig. 1d, e). Alternatively, multiple regression using smaller rival size and size difference as covariates showed that duration decreased with size difference but was not significantly affected by smaller rival size.

These analyses reveal a key difference in the relation between contest duration and larger rival (winner RHP) size depending on whether contests are resolved by mutual assessment or pure self-assessment. If contests are resolved by pure self-assessment, then contest duration should increase (weakly) with larger (winner) rival size in simple regression. In contrast, in cases of mutual assessment, duration should decrease strongly with larger (winner) rival size both in simple and in multiple regressions.

Taylor & Elwood (2003; summarized by Gammell & Hardy 2003) thus provide an analytical framework to distinguish the underlying mechanism driving animal contests. However, while 'pure self-assessment' (WOA-WA or E-WOA) can be identified, a problem arises in distinguishing CAM from mutual assessment. In the CAM there is no information about the opponent; however, costs accrue from the opponent's actions, and superior opponents are better at inflicting costs. This results in a 'true' negative relation between RHP difference and contest duration since, for any given RHP of the weaker rival, costs accrue faster, and weaker rival thresholds are reached sooner, as RHP of the stronger rival increases. Consequently, the CAM produces the same relation between stronger rival RHP and fight duration as predicted by mutual assessment. That is, the CAM will have the appearance of mutual gathering of information even though the decision is based on individual thresholds of cost. However, by staging contests between RHP-matched rivals (as recommended by Taylor & Elwood 2003) it should be possible to discriminate CAM from mutual assessment (Table 2). Pure self-assessment and CAM should both result in

Table 2
How to discriminate between assessment models

Experimental procedures	Pure self-assessment	CAM	Mutual assessment
Random pairings: relation between winner RHP and contest duration	Weak positive	Negative: of similar strength but opposite sign to that of weaker (loser) rival RHP and contest duration	Negative: of similar strength but opposite sign to that of weaker (loser) rival RHP and contest duration
RHP-matched pairings: relation between contestant RHP (typically size) and duration	Positive	Positive	No relation
Prior opponent observation	No difference in contest duration when fighting previewed or previously unseen opponent	No difference in contest duration when fighting previewed or previously unseen opponent	Shorter contests against previewed opponent compared to previously unseen opponent
Motivational probe technique	Negative relation between own RHP (size) and startle duration but no effect of opponent	Negative relation between own RHP (size) and startle duration but no effect of opponent	Negative relation between own RHP (size) and startle duration and positive relation between opponent RHP and startle duration

longer contests between strong RHP-matched rivals (e.g. Dixon & Cade 1986; Whitehouse 1997; Taylor et al. 2001), but this will not be seen with mutual assessment. With CAM, persistence should be positively related to size but costs inflicted should vary only with size difference. Thus with size-matched opponents duration is expected to increase with average size.

EVIDENCE FOR PURE SELF-ASSESSMENT

In the amphipod *Gammarus pulex*, single males (intruders) attempt to take over females held in precopula by other males (owners; Dick & Elwood 1990). Intruders attempt to grab the female and tug her away from the owner's grasp and larger males have an advantage over smaller males in making and resisting take-overs (Ward 1983), but owners have a strong positional advantage over intruders (Dick & Elwood 1990). Prenter et al. (2006) found a strong positive relation between loser size and duration and a nonsignificant positive relation between winner's weight and contest duration (Fig. 2), which indicates pure self-assessment (Taylor & Elwood 2003). The use and depletion of glucose reserves, as measured by the proportion of energy (glucose + glycogen) in the form of glucose, indicates physiological cost. For losers (intruders) this relation was negative and marginally nonsignificant, indicating fighting is costly and these physiological costs increase with loser weight. For winners (owners) this relation was nonsignificant and negative but less so than for losers (i.e. weakly negative), which is consistent with pure self-assessment (Prenter et al. 2006).

Similarly, in contests between male Cape dwarf chameleons, Stuart-Fox (2006) found that, in simple regressions, contest duration, intensity and the number of behavioural elements used (a measure of complexity) all showed a weak positive relation with winner ability, which supports pure self-assessment. Also, when winner RHP, loser RHP and a measure of contestant asymmetry were included as predictors of duration and intensity in multiple regressions, in each case loser RHP was the only variable retained in the final model after stepwise selection, which again is strong support for pure self-assessment. Stuart-Fox (2006) also examined those contests in which opponents were broadly matched and found that contest duration, intensity and number of behavioural
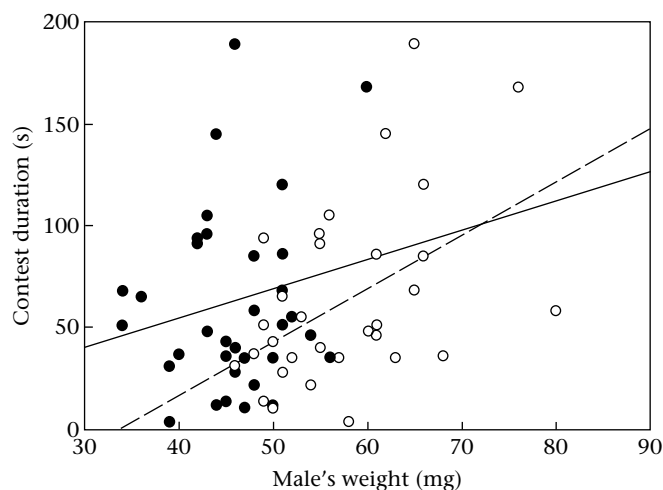
elements were all significantly associated with the mean ability of the two contestants, again consistent with pure self-assessment. In the fig wasp *Sycoscapter* sp. A, males also appeared to fight using a pure self-assessment strategy (Moore et al. 2008) because duration increased with loser's body size, but was unaffected by winner's size. In addition, the studies of spiders noted previously clearly point to pure self-assessment (Bridge et al. 2000; Taylor et al. 2001).

EVIDENCE FOR CAM

Morrell et al. (2005) observed male–male contests between burrow owners and intruders in the Australian fiddler crab *Uca mjoebergi*. Fighting was size assortative such that intruders tended to fight residents that were of a similar size to themselves. Two possibilities for this were suggested: (1) small individuals may be able to assess their relative inferiority prior to a fight with much larger individuals, and attempt to avoid the fight by retreating down the burrow (when the resident) or selecting a different opponent (when the intruder), and (2) large intruders may choose to avoid challenging residents much smaller than themselves since their small size would indicate a small burrow, which may be unsuitable (Jennions & Backwell 1996). Both possibilities suggest the decision to begin a fight may be based on an initial assessment of the potential opponent (Morrell et al. 2005).

In a multiple regression, using winner's and loser's sizes as independent predictors of duration, both factors remained significant, that is, duration increased with increasing size of the loser but decreased with increasing winner's size for a given size of loser (Morrell et al. 2005). However, according to Taylor & Elwood (2003), if only mutual assessment is occurring, the effect sizes should be approximately equal in magnitude and opposite in direction, but Morrell et al. (2005) found that the loser's size had a significantly stronger effect on duration than the winner's size. Simple regressions, investigating relations between measures of individual size and duration, showed the loser's size was the better
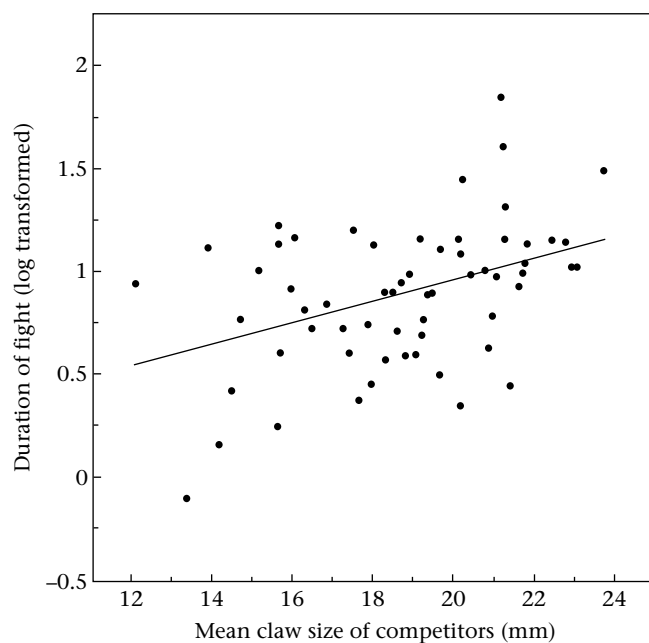
Figure 2. Relation between male weight and fight duration (s) for contests between adult male *Gammarus pulex* won by the defender. Black circles: winners (defenders); white circles: losers (attackers). The solid line represents the regression between winner's weight and contest duration ($R^2 = 0.005$, $P = 0.289$); the dashed line shows the regression between loser's weight and duration ($R^2 = 0.203$, $P = 0.005$). From Prenter et al. (2006), reprinted with permission of Elsevier.

Figure 3. Relation between the mean claw size of two size-matched competitors and fight duration in the fiddler crab *Uca mjoebergi*. The regression line is $Y = 0.054X - 0.107$ ($R^2 = 0.158$, $P = 0.001$). From Morrell et al. (2005), reprinted with permission of Elsevier.

predictor of contest duration, with the winner's size giving a weaker positive correlation with duration (as predicted by Taylor & Elwood 2003 for pure self-assessment). In addition, Morrell et al. (2005) found that in fights between closely size-matched individuals, fight duration increased with increasing mean size of the competitors (Fig. 3 as noted by others e.g. Dixon & Cade 1986; Glass & Huntingford 1988; Jennions & Backwell 1996; Whitehouse 1997; Taylor et al. 2001). This is consistent with pure self-assessment or CAM, but inconsistent with mutual assessment. Morrell et al. (2005) subsequently used a simulated, computer-generated population specifically to model the predictions of the CAM. To do this they used a model in which fighting was size assortative and persistence was based on individual size-determined cost thresholds, and opponents inflicted costs at a rate proportional to their size (i.e. larger rivals inflicted costs at a higher rate). Thus they defined the persistence duration of an individual as (own threshold − 1/3 × size of opponent). This simulation revealed similar results both in simple and in multiple regressions to the observed data, leading Morrell et al. (2005) to conclude that CAM was the most appropriate description of fighting behaviour in U. mjoebergi.

Recently, Briffa (2008) postulated the CAM as the most appropriate description of fighting between pairs of male house crickets, Acheta domesticus. Duration increased with loser's weight and decreased with winner's weight, as predicted by the assumption that contests are resolved by mutual assessment or CAM (Taylor & Elwood 2003). Briffa (2008) argued for the CAM, in which the loser gives up when the accumulated costs of fighting cross a threshold. Support for this came from the finding that individual thresholds of energetic cost contributed to the decision to give up. Namely, winners had higher glucose levels than losers and it is possible that low glucose level was related to the giving-up decision made by losers.

EVIDENCE FOR MUTUAL ASSESSMENT

Kemp et al. (2006) examined territorial aerial contests in the solitary wasp Hemipepsis ustulata. Males defend hilltop territories by means of elaborate noncontact repeat-ascending aerial interactions that may last for over an hour (Alcock & Bailey 1997). The protocol comprised three steps: (1) capture and retention of a known resident, (2) observation of a territorial replacement at the site for 30 min, and (3) release of the initial resident. This process ensures an escalated contest by eliminating the usual conventional residency effects that apply in this system. The authors presented several lines of evidence for mutual assessment or CAM during contests. They found contest duration was positively related to replacement (loser) size and negatively related to resident (winner) size, and these opposing relations were of roughly equal magnitude (Fig. 4). Also, in multiple regression, size asymmetry was significant, whereas absolute replacement size was not, again consistent with mutual assessment or CAM (Taylor & Elwood 2003). Further analysis of contests between size-matched opponents provides support for mutual assessment rather than CAM, as there was no relation between contestant size and duration (Table 2). In addition, since the interactions do not involve direct physical contact, it is unlikely that significant costs would be incurred from the opponent's actions, making the CAM unlikely to apply. Thus, the results of Kemp et al. (2006) provide strong evidence that mutual assessment occurs in contests between male H. ustulata.

A number of other studies also provide support for mutual assessment. For example, the single enlarged claw of male fiddler crabs, Uca pugilator, plays a major role in contests, which consist of a number of behavioural elements (Pratt et al. 2003). Fighting ability (RHP) is correlated with carapace width (a measure of body size) and the size of the claw (Hyatt & Salmon 1978; Jennions & Backwell 1996). Contest duration was positively correlated with the size category of the smaller contestant and negatively correlated with the difference in size between opponents (as predicted by all hypotheses; Taylor & Elwood 2003). However, when both size of the smaller competitor and size difference were modelled together, only the effect for size difference was significant. This provides support for both mutual assessment and CAM. However, when contestants were matched for size, size did not correlate with contest duration, which implies the underlying mechanism cannot be pure self-assessment or CAM, leaving mutual assessment as the better explanatory model. Similarly, Leiser et al. (2004) found that in contests between size-matched pairs of male
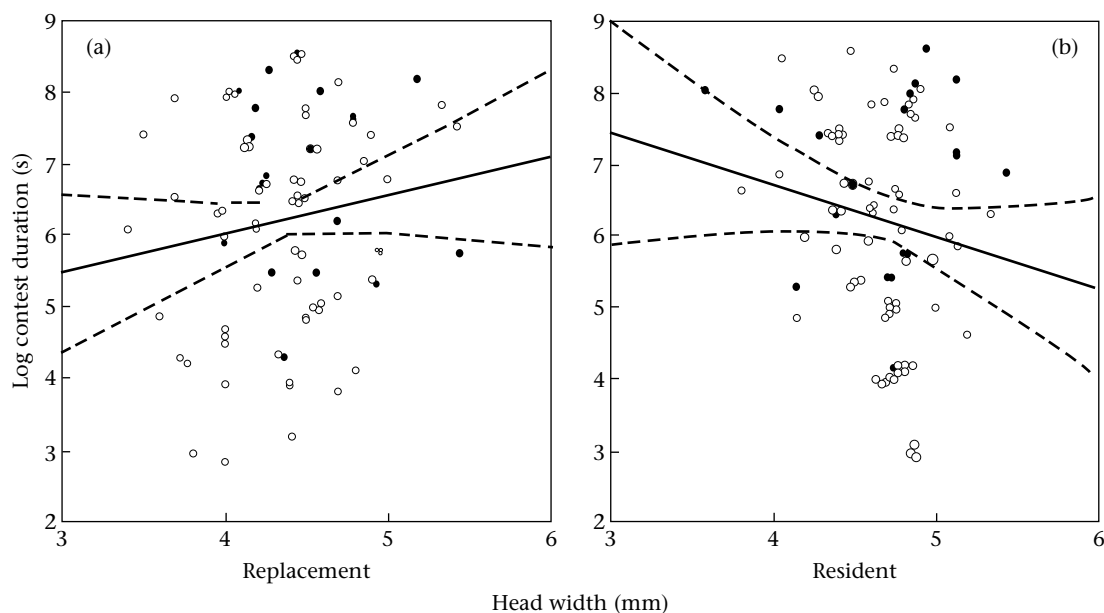

Figure 4. Linear relations (±95% confidence intervals) between contest duration and the head width of (a) replacement and (b) resident male *Hemipepsis ustulata* wasps. Solid circles indicate cases where replacement residents won. From Kemp et al. (2006), reprinted with permission of Elsevier.

convict cichlids, *Archocentrus nigrofasciatus*, large-pair and small-pair contests were of similar average duration, again supporting mutual assessment.

CONFLICTING AND INCONCLUSIVE RESULTS

In male–male contests in the Wellington tree weta, *Hemideina crassidens*, males with larger weapons (head length) have greater RHP and are more likely to win (Kelly 2006). Kelly (2006) found a negative relation between RHP asymmetry and both contest duration and intensity, as predicted by all three models of assessment (Taylor & Elwood 2003). There was a strong positive relation between smaller rival RHP (or loser RHP) and contest duration, which Kelly (2006) stated provides evidence for self-assessment, even though it is also predicted by mutual assessment and CAM. There was a negative trend for the relation between contest duration (and intensity) and larger male RHP (and winner RHP), which Kelly (2006) suggested indicates mutual assessment. However, since this relation did not reach significance and was not of similar strength to the positive relation of smaller rival (loser) RHP, the results provide no clear support for any model.

Jennings et al. (2004) investigated assessment between male fallow deer, *Dama dama*, during the rut. They used body mass and antler size as indicators of RHP, but, somewhat surprisingly, neither predicted victory (Jennings et al. 2004). Also, fight duration was not related to asymmetry in body weight or antler size, or to the weight or antler size of the lighter or heavier individual and thus the results appeared to show a lack of support for any model. Jennings et al. (2004) did, however, find some support for individuals basing escalation decisions on their own ability (self-assessment), because there was a positive relation between the loser's body weight and use of the jump clash to initiate antler contact. Thus losing animals increased their use of high-risk strategies as their body mass increased (similar to Bridge et al. 2000; Taylor et al. 2001). Further tentative support for self-assessment comes from the findings of a positive relation between the body weight of the lighter member of a dyad and the number of jump clashes recorded, and a nonsignificant positive relation between body weight of the heavier opponent and the number of jump clashes (i.e. similar to the predictions of Taylor & Elwood 2003 for pure self-assessment). However, the main predictions of the various models are not upheld.

Using a novel experimental approach, Rillich et al. (2007) provided evidence for opponent assessment, but not related to own ability in the Mediterranean cricket *Gryllus bimaculatus*. Pairs of contesting male crickets were subjected to various manipulations to investigate mechanisms underlying assessment. The mouthparts (mandibles and maxillae) appeared to be important external signals used for assessment, because when they were disabled fights lasted longer, were more intense and more often physical (Rillich et al. 2007). Also, visual cues were important and used in the assessment of opponent body size. In fights between small 'blinded' crickets and large normal crickets encounters lasted longer and were more aggressive than those between normal weight-mismatched crickets (Rillich et al. 2007). However, Rillich et al. (2007) argued that the decision to flee is based solely on assessment of the opponent's actions, which is contrary to mutual assessment in which a comparison of opponent with self is made to assess the relative advantages and disadvantages. In support of this, crickets lacking mandibles, and thus having a clear weaponry disadvantage, did not fight for significantly shorter periods or win significantly fewer fights against normal adversaries, compared to fights between two normal controls. Similarly, salamanders, *Plethodon cinereus*, with autotomized tails do not alter their own aggressive behaviour, although salamanders do become more aggressive towards opponents that have autotomized tails (Wise & Jaeger 1998). Rillich et al.

(2007) concluded that visual and physical cues act as external agonistic signals, which suppress the aggressiveness of the perceiver. Thus it is possible that, in addition to pure self-, CAM and mutual assessment, there may be a fourth model in which contestants assess opponent cues independent of their own ability or state. Few studies to date have taken such a proximate approach, manipulating cues potentially used for assessment (although see Briffa & Elwood 2000a; Tibbetts & Lindsay 2008). Future studies may benefit from a proximate approach, investigating the mechanisms used for assessment. However, consideration must be given to the ethical implications of invasively manipulating the sense organs of an animal, which may limit such studies.

ALTERNATIVE APPROACHES TO DISCRIMINATE MODELS

Use of Cues

A variety of studies have examined whether particular sensory cues from opponents influence the progress and outcome of contests. For example, Keeley & Grant (1993) provided a method to investigate whether visual cues are used to assess the opponent's fighting ability. They staged contests between pairs of convict cichlids that were matched for size and gender. Prior to the fight the contestants were separated by either a clear or opaque divider to allow or prevent visual contact, respectively. The results suggest opponents assess each other. Contests were shorter in the clear than in the opaque treatment, presumably because of the information acquired by observing and interacting with the opponent through the clear divider prior to the fight. Similarly, Earley et al. (2003), using green swordtails, *Xiphophorus helleri*, found contests were less escalated among males separated by a clear rather than an opaque divider beforehand. Similarly, rainbow trout, *Oncorhynchus mykiss*, settle conflicts with previewed opponents faster and with less aggression than conflicts with unfamiliar fish (Johnsson & Akerman 1998). Trout that lost contests against a previewed competitor reduced their aggression more rapidly than individuals that lost against a previously unseen opponent, while those that won against a previously observed competitor displayed a more rapid increase in aggression compared with winners that had faced an unfamiliar opponent (Johnsson & Akerman 1998). Furthermore, in pigs, *Sus scrofa*, Jensen & Yngvesson (1998) found that contests were significantly shorter between individuals that had been visually pre-exposed to one another compared to control pairs. Other pig studies, however, have found inconsistent results, for example Fraser (1974) found that prior exposure reduced aggression in pigs but Rushen (1988) failed to find any effect of prior familiarity on the incidence of fighting (although there was a nonsignificant trend for shorter fights between pre-exposed pairs of pigs). With the above studies, however, the possibility remains that the shorter observed contests in the clear (or pre-exposure) treatment could be the result of fatigue or energy depletion from attempts to interact with the opponent prior to the contest. Such a confounding variable could be overcome in future studies by having a control group in which the focal animal also previews a potential opponent but is then pitted against a different animal in the actual contest. This approach was recently adopted in a study by delBarco-Trillo et al. (2009), using male Turkish hamsters, *Mesocricetus brandti*. They housed pairs of weight-matched male hamsters in the same cage for 48 h, separating each individual of the pair by a wire-mesh barrier, enabling visual, auditory, olfactory and tactile stimulation between males, but preventing fighting. Subsequently, after 48 h cohabitation, males were pitted against either the familiar cagemate or an unfamiliar opponent from another pairing. In line with predictions for assessment, pre-exposure to an opponent reduced aggression, resulting in a decrease in contest duration, the number of fights and

the percentage of time spent fighting, while the latency to engage in a fight increased with familiarity (delBarco-Trillo et al. 2009).

Visual assessment of strength has recently been demonstrated in humans (Sell et al. 2009). First, the strength of males was measured by their ability to lift weights. Subsequently, subjects were asked to rate the strength and perceived toughness (a measure of fighting ability) by viewing photographs of those males. Perhaps unsurprisingly given our highly evolved cognitive abilities, together with anthropological evidence indicating the importance of aggression among our ancestors (e.g. Manson & Wrangham 1991), subjects were able to rate the strength accurately from photographs of the full person, the body alone, and just the face. Perhaps our own aptitude for mutual assessment may have led us generally to assume such capabilities in other species. Similarly, rival assessment, based on a badge of status, was clearly demonstrated in female *Polistes dominulus* wasps (Tibbetts & Lindsay 2008). These wasps have a conspicuous black facial pattern that varies between individuals and is associated with dominance (Tibbetts & Dale 2004) and condition (Tibbetts & Curtis 2007). The greater the disruption or 'brokenness' of the black facial pattern, the higher is the wasp's quality. Focal wasps were given access to two food patches, each guarded by a wasp, the facial pattern of which had been experimentally manipulated. Subsequently, wasps chose to challenge guards with facial patterns indicating a low level of quality, avoiding those with more facial 'spots' indicative of a high level of quality. Furthermore, during agonistic territorial interactions between male lizards, visual cues have been shown to influence contest progression and outcome. For example, Huyghe et al. (2005) found that in *Gallotia galloti* lizards, winners tended to have larger total areas of blue patches on their sides, suggesting that these badges convey information on male social status. Similarly, dewlap displays, in which the extendable throat fan is flared at an opponent during contests, are thought to provide visual information on fighting ability (Jenssen et al. 2000; Vanhooydonck et al. 2005), particularly since dewlap size has been shown to correlate with determinants of RHP, such as bite force and body size, in the green anole lizard, *Anolis carolinensis* (Vanhooydonck et al. 2005), although experiments manipulating dewlap size, as yet, remain to be conducted and are necessary to reveal if, indeed, the dewlap is used in assessment.

Auditory cues have also been shown to influence fight decisions. For example, by experimentally manipulating the ability of male toads, *Bufo bufo*, to vocalize during contests for the possession of females, Davies & Halliday (1978) showed the importance of calls as auditory signals. Call frequency is related to body size in toads, with larger males having deeper croaks. Davies & Halliday (1978) silenced males in possession of a female and subsequently played either the croak of a small male or that of a large male to potential attackers. Subsequently, the attacker's behaviour depended on the call frequency it heard, with attack being much more likely when a high-pitched croak, indicative of a small male, was played to a potential attacker. With the relation between call frequency and body size and therefore RHP, this signal appears to act as a reliable indicator of fighting ability that is assessed by an opponent. Similarly, roaring contests and parallel walk displays have long been postulated as auditory and visual cues used for opponent assessment during red deer, *Cervus elaphus*, stag contests (e.g. Clutton-Brock & Albon 1979).

In terms of tactile cues, the temporal pattern and power of shell rapping during hermit crab, *Pagurus bernhardus*, contests transmits information relating to the RHP of the attacker and influences the decision of the defender to resist or allow itself to be evicted (Briffa et al. 1998, 2003; Briffa & Elwood 2000a, b, c). Attackers that had their shell coated with rubber, thus reducing the impact of their raps, were less likely to evict defenders, showing that the power

(and possibly the pattern) of raps is perceived and used in fight decisions (Briffa & Elwood 2000b).

While assessment of visual, auditory and tactile cues of RHP has arguably received the most attention, other sensory modalities have been demonstrated. For example, intriguingly, electric signals produced by fish may act as important cues providing information during contests (Triefenbach & Zakon 2008). In staged dyadic contests over a shelter, male knifefish, *Apteronotus leptorhynchus*, modulated the frequency of their electric organ discharge. Through the course of an agonistic encounter, ultimate winners emitted significantly more high-frequency electrical signals termed 'chirps', suggesting these cues may act as conventional signals providing the opponent with reliable information on aggressive motivation and fighting ability (Triefenbach & Zakon 2008).

Chemical cues are another source of potential information available to contestants, and detection of chemical signals during aggressive encounters through the use of antennae and antennules has been well documented in crustaceans (Bergman et al. 2003). In American lobsters, *Homarus americanus* (Karavanich & Atema 1998), and Norway lobsters, *Nephrops norvegicus* (Katoh et al. 2008), chemical cues are assessed and influence the progression and outcome of agonistic encounters. Chemical cues are also assessed during crayfish, *Orconectes rusticus*, contests (Zulandt-Schneider et al. 2001; Bergman et al. 2003). When urine release was blocked in adult males, fights were significantly longer and reached a higher intensity level, indicating that under normal circumstances information is transmitted through chemicals released with the urine, influencing the behaviour of an opponent (Zulandt-Schneider et al. 2001). Furthermore, Bergman et al. (2003) demonstrated assessment of odour signals mediating a winner effect in crayfish. Taken together, the experimental manipulations used in these studies on crustacea reveal the importance of urine-carried chemical signals, perceived by the antennules, and assessed to mediate the decisions an individual makes during an aggressive encounter. When deprived of such cues, an individual has reduced ability to determine the status of its opponent, and may inappropriately fight more intensely and for longer periods.

The importance of chemical communication has been further examined during parasitoid wasp contests using an ingenious experimental set-up (e.g. Goubault et al. 2006, 2008). In these studies, it was possible to follow a 'chemical conversation', by analysing emissions from contestants in real time using a mass spectrometer linked to the contest arena. Furthermore, to enable the chemical emissions from each individual of an aggressively interacting pair to be tracked independently during contests, a technique was used to manipulate the molecular mass of the chemical signals of particular opponents by rearing some parasitoids on deuterium-enhanced hosts. In this way, by staging contests between a deuterated and an undeuterated individual, it was possible to track the chemical emissions from each contestant. Goubault et al. (2006) staged contests between *Goniozus legneri* females over paralysed hosts. Females losing an agonistic encounter were found to emit a volatile spiroacetal compound from the head (also see Goubault et al. 2008). Future studies may seek to use this novel technique when investigating chemical assessment, since it has the potential to relate chemical release to changes in behaviour consistent with the opponent having gathered information from the emission.

The modification of behaviour as a result of information gathered from opponent cues is consistent with some form of mutual assessment; however, care must be taken with this interpretation. For example, despite the presence of apparent cues relating to fighting ability, the opponent may still fail to gather information (pure self-assessment and CAM). Furthermore, we may measure fighting ability using one particular character, such as body size; however, this is only a correlate of RHP, and it is possible that an

animal might assess a different cue (if mutual assessment occurs) or not assess at all. Furthermore, we may incorrectly assume that opponents gather information about RHP from particular fight activities that are obvious to us. Thus, we see spiders use visual and vibratory signals, yet information about RHP appears not to be gathered (e.g. Bridge et al. 2000; Taylor et al. 2001). In contests between male *G. pulex* amphipods, the intruder pulls and tugs at the owner male, and female held in precopula, offering potential tactile cues, but these appear not to be assessed (Prenter et al. 2006). This highlights the need for experimental approaches, manipulating the signal in question, such as the dewlap of lizards, to determine any subsequent changes in opponent behaviour, indicative of some form of assessment.

Probing Motivation

Game theory predicts that information gathering about the opponent and/or own abilities should alter the motivational state of each opponent in a contest (Parker & Stuart 1976). The motivational state of an animal may be probed by the use of a novel, potentially startling stimulus, the speed of recovery reflecting the motivation to continue the previous activity (e.g. Culshaw & Broom 1980), and the technique has been applied to fighting animals (Elwood et al. 1998). The problem is to startle just one of an interacting pair. In hermit crab shell fights one crab withdraws into the shell but the other (attacker) is partially emerged and open to being startled (Elwood et al. 1998; Briffa & Elwood 2001b). However, for other animal contests, where both opponents have a similar fighting method, a technique is required so that just one opponent is subject to the novel startling stimulus. This has been developed for fish, using a method in which a novel stimulus is applied to cause a startle response in one contestant occupying a clear-sided tank, and displaying to an opponent in an adjacent tank, by dropping a glass marble in such a way that it is only perceived by one fish (G. Arnott & R. Elwood, unpublished data). Startle duration can then be used to infer any information the startled fish has about its opponent's and/or its own RHP. Furthermore, in a contrived situation in which the opponents cannot inflict direct costs, for example, in staged agonistic encounters between subjects aggressively displaying at each other from separate tanks, or separated by a clear partition, pure self-assessment and CAM should have the same predictions, meaning mutual assessment can clearly be discriminated from CAM. With pure self-assessment and CAM, startle duration should be significantly negatively related to the 'startled' individual's weight (RHP), indicating increased motivation to fight as own ability increases, but there should be no relation between the startle and the opponent's weight (RHP; Fig. 5a). By contrast, with mutual assessment, the negative relation between the 'startled' individual's weight and startle duration should remain but there should also be a significant positive relation between the 'nonstartled' opponent's weight and startle duration, indicating decreased fight motivation of the 'startled' individual as opponent ability increases (Fig. 5b). Such an approach may provide a useful tool for future studies investigating assessment abilities. Furthermore, the probe may be applied at any point in a contest enabling us to enquire what information has been gathered at that time and thus there is the potential to plot motivational change (Briffa & Elwood 2001b). If used relatively early in an aggressive interaction it negates the need for contestants to engage in potentially lengthy, welfare-compromising contests (Huntingford 1984).

VARYING DEGREES OF ASSESSMENT?

Prenter et al. (2006) suggested that pure self-assessment and mutual assessment are extremes of a continuum of possible
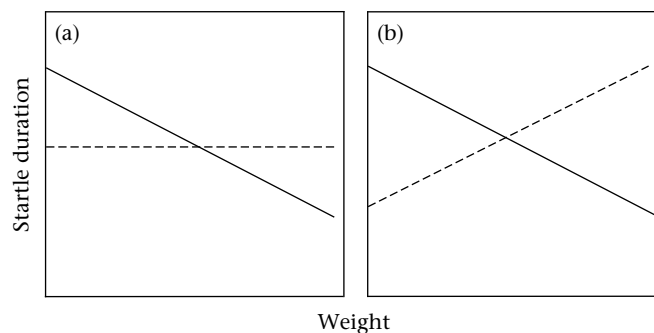

Figure 5. Relations when 'focal' weight and 'stimulus' weight are used as independent predictors of 'focal' startle duration (log) in multiple regression. (a) Pure self-assessment (and CAM): negative relation between 'focal' fish weight and startle duration, but no relation for 'stimulus' fish weight. (b) Mutual assessment: negative relation between 'focal' weight and startle duration and positive relation between 'stimulus' fish weight and startle duration. Solid line represents 'focal' fish, dashed line represents 'stimulus' fish.

assessment strategies in animal contests rather than alternatives. It is likely that an animal will always have good knowledge of its own abilities. If it is able to gather equally reliable information about its opponent we have mutual assessment. If it is not able to gather any information about the opponent we have pure self-assessment. However, an animal might be able to gather some information about the opponent but not enough to equal that about itself (Prenter et al. 2006). Thus some contests could produce a situation where there is 'partial mutual assessment' (Prenter et al. 2006). This should still produce a strong positive relation between size (RHP) and cost to the loser, but the relation for winners may shift from a weak positive relation between size and cost to a neutral relation and then to a negative relation as increasing information about the opponent is possible (Fig. 6; Prenter et al. 2006).

This suggestion by Prenter et al. (2006) is further supported by recent work on assessment in shore crabs, *Carcinus maenas* (Smallegange et al. 2007) and jumping spiders, *Phidippus clarus* (Elias et al. 2008), while the results of Kelly (2006) for Wellington tree weta contests are also consistent with the idea of 'partial mutual assessment'. Smallegange et al. (2007) staged fights between pairs of crabs and used a statistical approach similar to that recommended by Taylor & Elwood (2003). The results were somewhat inconclusive. Both size of the smaller competitor and absolute size difference affected contest duration, but this does not distinguish between models. Further analysis confirmed the (positive) effect of the size of the smaller competitor on contest duration, but also showed a trend that the size of the larger
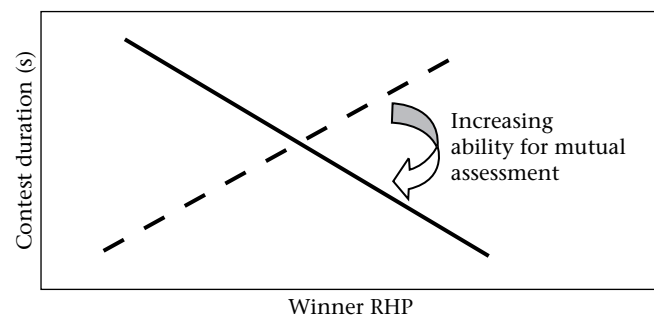

Figure 6. Representation of the hypothesized relation between contest duration and the winner's RHP with varying degrees of information gathered about the opponent. The dashed line represents the relation with self-assessment and the solid line shows the expected relation with accurate mutual assessment. From Prenter et al. (2006), reprinted with permission of Elsevier.

competitor affected contest duration in a negative manner, hinting at mutual assessment or CAM (Smallegange et al. 2007). Additionally, the duration of contests between size-matched contestants was constant regardless of the absolute size of the crabs, which also points to mutual assessment. These inconsistent findings led Smallegange et al. (2007) to suggest that shore crab contests may involve a situation analogous to that of 'partial mutual assessment' suggested by Prenter et al. (2006). Namely, a crab is likely to have good knowledge of its own competitive ability. During a contest it may gather information on the opponent's competitive ability, which is not as reliable as the information on its own ability. Such a scenario would produce a weak negative relation between the RHP of the winner and contest duration (Prenter et al. 2006), which is consistent with the nonsignificant trend that, as the size of the larger crab (usually the winner) increased, contest duration decreased. Further evidence comes from male jumping spider, P. clarus, contests over access to females: Elias et al. (2008) found that when loser and winner weights were considered separately, in simple linear regressions, loser weight was significantly positively related to contact duration, whereas winner weight showed a nonsignificant tendency to relate negatively to contact phase duration, which is consistent with a scenario of partial mutual assessment (Prenter et al. 2006). However, it is likely that distinguishing cases of partial mutual assessment from pure self-assessment will be difficult.

We may also get the same variation in the relation of the larger animal (winner) and contest duration with the CAM. In this model the threshold of costs is set by the animal's own RHP but how quickly that is reached will depend on two factors: first, the actions that the animal makes that are independent of what the opponent does to it, and second, the effects of what the opponent does to it. If the latter are minor the animal is essentially fighting according to self-assessment rather like the WOA-WA (Mesterton-Gibbons et al. 1996) and E-WOA (Payne & Pagel 1996, 1997). However, if the effects of the opponent are large then regression analysis will appear as if there is mutual assessment. Thus, similar to differing amounts of information being gathered (Fig. 6), the relation between the RHP of the winner and contest duration will swing from slightly positive to strongly negative as the effects of the opponent increase, even though CAM may be the best explanatory model (e.g. Morrell et al. 2005). In those cases in which CAM is appropriate but the effect of the opponent is minor, compared to those of self, it will be difficult to distinguish from pure self-assessment.

SWITCHING ASSESSMENT STRATEGIES?

It is possible that different assessment mechanisms are used during different phases of a contest. For example, the findings of Morrell et al. (2005) on size-assortative fighting suggest that some assessment of the opponent may occur before an individual decides to engage in an interaction, although later they appear to fight according to CAM. Similarly, in the hermit crab P. bernhardus (Elwood et al. 2006), attacking animals choose to initiate encounters with opponents smaller than themselves suggesting opponent assessment. However, they then appear to switch assessment strategy as described below. Thus assessment of opponents could occur during one phase of a contest, while individual thresholds may be important in another, a possibility also highlighted by Stuart-Fox (2006).

To investigate whether different assessment mechanisms are used during different phases, Morrell et al. (2005) suggested that data on the separate phases of the fight (e.g. the push and grapple phases in fiddler crab fights) would be needed. Recently, Hsu et al. (2008) adopted such an approach and found evidence for switching

assessment strategy during killifish, Kryptolebias marmoratus, contests. These usually begin with mutual displays involving erected gill covers and may then proceed to an attack by one opponent and escalated fighting if the other contestant fights back. Two types of contest were staged, random-sized pairings and equal-sized pairings. In addition to testing the predictions of Taylor & Elwood (2003) for overall contest duration, Hsu et al. (2008) also analysed separately those contests resolved after mutual display, after one attack, or after an escalated fight.

For the random-sized pairing contests, total contest duration tended to increase with the size of the smaller opponent and decrease with the size of the larger opponent, consistent with predictions of mutual assessment or CAM. For the size-matched pairings, there was no relation between pair size and contest duration, consistent with mutual assessment but not CAM. However, when contests resolved at different stages were analysed separately, a more complex scenario emerged. The likelihood that a contest was resolved with mutual displays was related positively to the size of the larger opponent (as the size of the larger opponent increased, the eventual losers were more likely to give up before any attacks occurred), and showed a trend to relate negatively to the size of the smaller opponent, thus pointing towards mutual assessment (CAM is unlikely to apply to a display phase). Furthermore, the likelihood of contest resolution after one attack was related negatively to the size of the attack receiver and positively to the size of the attacker, again consistent with mutual assessment or CAM. Finally, for escalated contests, the duration of escalation showed a strong positive relation to the size of the loser and a weak positive (nonsignificant) relation to that of the winner, consistent with pure self-assessment. Thus, K. marmoratus individuals appear to use information about their opponent, or to be influenced by the opponent, early in a contest. However, once the contest is escalated, their opponent's abilities no longer influence their fighting decision, and they persist in accord only with their own threshold of costs (Hsu et al. 2008).

This influential paper by Hsu et al. (2008) highlights the benefit of examining different phases of the contest separately. Had they only used contest duration to test the assessment strategy used, they would have concluded the fish adopt mutual assessment throughout the contest. However, they switch from early mutual assessment to later self-assessment, which is contrary to the prediction of SAM (Enquist & Leimar 1983; Enquist et al. 1990) whereby the quality of opponent assessment increases as the contest proceeds.

More recently, Moore et al. (2008) also tested for switching assessment strategies in fig wasps. One species appeared to use a prefight assessment strategy, with body size and mandible length differences between fighting male pairs being larger than those between randomly chosen pairs from figs, leading the authors to suggest males were assessing and attacking inferior rivals to remove them from the competitor pool. Subsequently, during the escalated fights attacking males appeared to switch to a self-assessment strategy.

DIFFERING ASSESSMENT BETWEEN CONTESTANTS

It is not unusual for opponents to adopt two distinct roles, with different fight activities, particularly when the contest involves one individual attempting to acquire a resource in the possession of its opponent (Briffa & Elwood 2004). Examples include contests over guarded females in newts (Verrell 1986), amphipods (Dick & Elwood 1990) and damselflies (Cordero 1999). This raises the question, are there circumstances where opponents differ in the assessment strategy used?

This is the case in the hermit crab *P. bernhardus*. In contests for shells, the opponents adopt two very distinct roles. In a typical interaction, fights are initiated by the larger of the two contestants, termed the 'attacker', against the 'defender'. The attacker grabs the defender's shell and, after assessing the potential gain (Elwood et al. 1998), may escalate the fight to 'shell rapping', in which the attacker hits its shell upon that of the defender, until either the defender is evicted, enabling the attacker to take that shell, or the attacker gives up. The defender, in contrast, remains tightly withdrawn into its shell until the encounter is resolved. The two roles differ with respect to the amount of information each contestant has about resource value (Arnott & Elwood 2007) and in the information they can gather about the opponent.

The defender clearly gains information about the attacker because defenders give up if attacked vigorously (Briffa et al. 1998, 2003; Briffa & Elwood 2000a, b, c). The duration of pauses between bouts of rapping is a predictor of attacker success; when the pauses are short, eviction of the defender is likely (Briffa et al. 1998). Defenders appear to assess the performance of attackers, and make a decision regarding whether to mount a defence of the shell or not and the former involves physiological change. Defenders that successfully defend their shells have greater concentrations of blood glucose than either those that have not fought or those that are evicted (Briffa & Elwood 2001a, 2002). Furthermore, there is no difference between glycogen concentration in evicted defenders and those that did not fight, indicating that the energy reserves of losers had not been mobilized (Briffa & Elwood 2004). Briffa & Elwood (2004) also found a positive relation between the proportion of the total energy pool (sum of circulating glucose and stored glycogen) in the form of glucose in defenders and the duration of attacker pauses between bouts of rapping. Defenders facing a weak attacker tend to attempt to resist eviction, with the mobilization of glycogen stores, leading to the elevation of blood glucose whereas this is not seen in those that give up. Thus defenders gather information about the attacker.

In contrast, defenders do not perform any overt activities that attackers could assess (Briffa & Elwood 2005). Rather, attackers seem to use a self-assessment strategy, based only on their own physiological state, fighting up to a threshold of fatigue, shown by lactic acid, which accumulates as the fight progresses (Briffa & Elwood 2001a, 2002). In attackers, those that give up have high lactate (Briffa & Elwood 2001a) and lactate increases with the number of bouts of rapping (Briffa & Elwood 2005). Lactate constrains performance and is used in the attacker's decision making.

Further support suggestive of asymmetry in information use between contestants has recently been observed in male swordtail, *X. helleri,* contests (Prenter et al. 2008). Initially, the results appeared to point towards mutual assessment, with the finding that in size-matched contests duration did not increase with mean body size. However, the picture was somewhat more complex. The 'sword' in these fish comprises a set of multi-coloured fin rays extending well beyond the caudal margin (Benson & Basolo 2006), and large swords confer an advantage in attracting females (Basolo 1990). Prenter et al. (2008) found the sword of the winner had a marked effect on the decision of the loser to give up, with shorter fights when the winner had a long sword relative to its body size. Losers thus assessed the swords of winners, which precludes self-assessment. However, mutual assessment predicts that the sword of the loser and relative sword length should have effects but this was not found. Furthermore, in a demonstration of differing assessment between opponents, there was no indication that the winner gathered information about the loser's sword. Thus the two opponents differ in their assessments (Prenter et al. 2008).

CONCLUSION

Examples noted above show that by judicious use of different techniques it is possible to discriminate between the three main models of assessment that differ in information gathering (Table 2). The relation between winner RHP and contest duration is strongly negative with CAM and mutual assessment but not with pure self-assessment. With matched pairs, contest duration will increase with pure self-assessment and CAM but not mutual assessment. Prior observation of an opponent should result in swifter resolution in cases of mutual assessment but not with CAM or pure self-assessment, but care must be taken to have a control in which the focal animal sees another potential opponent. Finally, the motivational probe should show effects of both own size and opponent size in cases of mutual assessment but not with CAM or pure self-assessment. Furthermore, examining durations of different contest stages may also indicate whether the information-gathering abilities or use of information changes from stage to stage. Thus, future studies need to apply a battery of techniques to determine how animals settle contests.

However, a number of studies examined here (e.g. Morrell et al. 2005; Kelly 2006; Stuart-Fox 2006) have determined the assessment mechanism in terms of information as recommended above but then tested other specific factors relating to particular game theory models. For example, all WOA models predict that behaviour should be matched between contestants until near the end of the contest, so behavioural matching has been used to discriminate WOA models from both the CAM and SAM in which behaviour of opponents can be unmatched (Stuart-Fox 2006). Furthermore, to distinguish CAM from SAM an examination of escalation rates has been recommended (e.g. Kelly 2006; Stuart-Fox 2006). In the SAM, contests occur in discrete phases (periods of the contest characterized by activities of similar intensity) with interphase escalation, that is, each phase is more intense than the previous one (Enquist & Leimar 1983; Enquist et al. 1990), whereas in the CAM contests occur in a single phase and escalate (Payne 1998).

Such an approach, however, has led to conflict in assigning the most applicable model. For example, a number of studies detailed here found evidence for pure self-assessment, namely a positive relation between loser RHP and contest cost and a weaker positive relation between winner RHP and cost (e.g. Morrell et al. 2005; Prenter et al. 2006; Stuart-Fox 2006). Subsequently, Stuart-Fox (2006) detailed evidence in support of the CAM being the most applicable game theory model, while Prenter et al. (2006) hinted towards this being the case. For example, in the chameleon contests studied by Stuart-Fox (2006), opponents appeared not to match fighting intensity, thus seemingly ruling out the WOA models, and there was a temporal overlap of activities within contests, consistent with the CAM, which allows a gradual change in the proportion of activities used (Payne 1998). Similarly, Kelly (2006) argued that tree weta contests occur in a single phase and escalate, which is consistent with the CAM but not the SAM. However, as previously explained, while the CAM is based on self-assessment, it is also a game of endurance and tolerance to damage inflicted by an opponent (Payne 1998) and the decision to withdraw is influenced by both an individual's own RHP (poor-quality individuals can bear fewer costs) and the opponent's RHP (higher quality individuals can inflict costs at a higher rate; Payne 1998). Thus, clear support for pure self-assessment by use of the methods proposed by Taylor & Elwood (2003) precludes CAM as a possible model and subsequently suggesting support for that model is unwarranted (e.g. Prenter et al. 2006; Stuart-Fox 2006). In contrast, Rillich et al. (2007) detailed evidence supporting opponent assessment in Mediterranean cricket contests, but then argued support for CAM, which is surprising, given that CAM does not involve opponent

assessment. Furthermore, Briffa (2008) used the approach advocated by Taylor & Elwood (2003) to find support for CAM in male house cricket contests but subsequently questioned this support based on studies (Hack 1997a, b) showing both escalation and de-escalation, in terms of the choice of agonistic tactic, which is contrary to predictions for both CAM and SAM, whereby the intensity of agonistic behaviour should escalate as the fight progresses.

We have concerns regarding the use of specific criteria such as matching of behaviour and rates of escalation subsequently to accept or reject a particular model. Such observations are somewhat subjective and open to interpretation. For example, deciding upon whether contests are viewed as occurring in one phase (as predicted by CAM and E-WOA) or discrete phases (as predicted by SAM) is problematic and hinges on the degree to which elements are mixed across phase categories within a contest sequence. Furthermore, these models have been constructed within certain parameters, presumably for mathematical practicality, and thus may not always reflect reality. Therefore, the apparent fixation with fitting empirical findings to theoretical models may actually be hindering our understanding of information gathering and assessment of RHP.

Furthermore, the concepts of 'partial mutual assessment' (Prenter et al. 2006), 'switching assessment' (Hsu et al. 2008), 'differing assessment between contestants' (Briffa & Elwood 2004) and 'opponent only assessment' (Rillich et al. 2007), point to limitations of current game theory models. Thus, while acknowledging the importance of current models in shaping and enhancing our understanding of animal contests, we challenge biological game theoreticians to develop existing models, addressing the concerns raised above.

Acknowledgments

References

Adams, E. S. & Caldwell, R. L. 1990. Deceptive communication in asymmetric fights of the stomatopod crustacean *Gonodactylus bredini*. *Animal Behaviour*, **39**, 706–716.

Alcock, J. & Bailey, W. J. 1997. Success in territorial defence by male tarantula hawk wasps *Hemipepsis ustulata*: the role of residency. *Ecological Entomology*, **22**, 377–383.

Archer, J. 1988. *The Behavioural Biology of Aggression*. Cambridge: Cambridge University Press.

Arnott, G. & Elwood, R. W. 2007. Fighting for shells: how private information about resource value changes hermit crab pre-fight displays and escalated fight behaviour. *Proceedings of the Royal Society of London, Series B*, **274**, 3011–3017.

Arnott, G. & Elwood, R. W. 2008. Information gathering and decision making about resource value in animal contests. *Animal Behaviour*, **76**, 529–542.

Austad, S. N. 1983. A game theoretical interpretation of male combat in the bowl and doily spider (*Frontinella pyramitela*). *Animal Behaviour*, **31**, 59–73.

Barki, A., Harpaz, S. & Karplus, I. 1997. Contradictory asymmetries in body and weapon size, and assessment in fighting male prawns, *Macrobrachium rosenbergii*. *Aggressive Behavior*, **23**, 81–91.

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

Benson, K. E. & Basolo, A. L. 2006. Male–male competition and the sword in male swordtails, *Xiphophorus helleri*. *Animal Behaviour*, **71**, 129–134.

Bergman, D. A., Kozlowski, C. P., McIntyre, J. C., Huber, R., Daws, A. G. & Moore, P. A. 2003. Temporal dynamics and communication of winner-effects in the crayfish, *Orconectes rusticus*. *Behaviour*, **140**, 805–825.

Bridge, A. P., Elwood, R. W. & Dick, J. T. A. 2000. Imperfect assessment and limited information preclude optimal strategies in male–male fights in the orb-weaving spider *Metellina mengei*. *Proceedings of the Royal Society of London, Series B*, **267**, 273–279.

Briffa, M. 2008. Decisions during fights in the house cricket, *Acheta domesticus*: mutual or self assessment of energy, weapons and size? *Animal Behaviour*, **75**, 1053–1062.

Briffa, M. & Elwood, R. W. 2000a. Cumulative or sequential assessment during hermit crab shell fights: effects of oxygen on decision rules. *Proceedings of the Royal Society of London, Series B*, **267**, 2445–2452.

Briffa, M. & Elwood, R. W. 2000b. The power of shell rapping influences rates of eviction in hermit crabs. *Behavioral Ecology*, **11**, 288–293.

Briffa, M. & Elwood, R. W. 2000c. Analysis of the finescale timing of repeated signals: does shell rapping in hermit crabs signal stamina? *Animal Behaviour*, **59**, 159–165.

Briffa, M. & Elwood, R. W. 2001a. Decision rules, energy metabolism and vigor of hermit crab fights. *Proceedings of the Royal Society of London, Series B*, **268**, 1841–1848.

Briffa, M. & Elwood, R. W. 2001b. Motivational change during shell fights in the hermit crab *Pagurus bernhardus*. *Animal Behaviour*, **62**, 505–510.

Briffa, M. & Elwood, R. W. 2002. Power of shell-rapping signals influences physiological costs and subsequent decisions during hermit crab fights. *Proceedings of the Royal Society of London, Series B*, **269**, 2331–2336.

Briffa, M. & Elwood, R. W. 2004. Use of energy reserves in fighting hermit crabs. *Proceedings of the Royal Society of London, Series B*, **271**, 373–379.

Briffa, M. & Elwood, R. W. 2005. Rapid change in energy status in fighting animals: causes and effects of strategic decisions. *Animal Behaviour*, **70**, 119–124.

Briffa, M., Elwood, R. W. & Dick, J. T. A. 1998. Analysis of repeated signals during shell fights in the hermit crab *Pagurus bernhardus*. *Proceedings of the Royal Society of London, Series B*, **265**, 1467–1474.

Briffa, M., Elwood, R. W. & Russ, J. M. 2003. Change in the power of rapping during shell fights between hermit crabs: does magnitude contribute to repeated signals? *Behavioral Ecology*, **14**, 60–65.

Clutton-Brock, T. H. & Albon, S. D. 1979. The roaring of red deer and the evolution of honest advertisement. *Behaviour*, **69**, 145–169.

Clutton-Brock, T. H., Albon, S. D., Gibson, R. M. & Guinness, F. E. 1979. The logical stag: adaptive aspects of fighting in red deer (*Cervus elaphus* L). *Animal Behaviour*, **27**, 211–225.

Cordero, A. 1999. Forced copulations and female contact guarding at a high male density in a calopterygid damselfly. *Journal of Insect Behavior*, **12**, 27–37.

Culshaw, A. D. & Broom, D. M. 1980. The imminence of behavioural change and startle responses of chicks. *Behaviour*, **73**, 64–76.

Dale, S. & Slagsvold, T. 1995. Female contests for nest sites and mates in the pied flycatcher *Ficedula hypoleuca*. *Ethology*, **99**, 209–222.

Davies, N. B. & Halliday, T. R. 1978. Deep croaks and fighting assessment in toads *Bufo bufo*. *Nature*, **274**, 683–685.

delBarco-Trillo, J., McPhee, M. E. & Johnston, R. E. 2009. Nonagonistic familiarity decreases aggression in male Turkish hamsters, *Mesocricetus brandti*. *Animal Behaviour*, **77**, 389–393.

Dick, J. T. A. & Elwood, R. W. 1990. Symmetrical assessment of female quality by male *Gammarus pulex* (amphipoda) during struggles over precopula females. *Animal Behaviour*, **40**, 877–883.

Dixon, K. A. & Cade, W. H. 1986. Some factors influencing male–male aggression in the field cricket *Gryllus integer* (time of day, age, weight and sexual maturity). *Animal Behaviour*, **34**, 340–346.

Dowds, B. M. & Elwood, R. W. 1985. Shell wars 2: the influence of relative size on decisions made during hermit crab shell fights. *Animal Behaviour*, **33**, 649–656.

Draud, M. & Lynch, P. A. E. 2002. Asymmetric contests for breeding sites between monogamous pairs of convict cichlids (*Archocentrus nigrofasciatum*, Cichlidae): pair experience pays. *Behaviour*, **139**, 861–873.

Dugatkin, L. A. & Biederman, L. 1991. Balancing asymmetries in resource holding power and resource value in the pumpkinseed sunfish. *Animal Behaviour*, **42**, 691–692.

Dugatkin, L. A. & Ohlsen, S. R. 1990. Contrasting asymmetries in value expectation and resource holding power: effects on attack behaviour and dominance in the pumpkinseed sunfish, *Lepomis gibbosus*. *Animal Behaviour*, **39**, 802–804.

Earley, R. L., Tinsley, M. & Dugatkin, L. A. 2003. To see or not to see: does previewing a future opponent affect the contest behavior of green swordtail males (*Xiphophorus helleri*)? *Naturwissenschaften*, **90**, 226–230.

Elias, D. O., Kasumovic, M. M., Punzalan, D., Andrade, M. C. B. & Mason, A. C. 2008. Assessment during aggressive contests between male jumping spiders. *Animal Behaviour*, **76**, 901–910.

Elwood, R. W., Wood, K. E., Gallagher, M. B. & Dick, J. T. A. 1998. Probing motivational state during agonistic encounters in animals. *Nature*, **393**, 66–68.

Elwood, R. W., Pothanikat, R. M. E. & Briffa, M. 2006. Honest and dishonest displays, motivational state and subsequent decisions in hermit crab shell fights. *Animal Behaviour*, **72**, 853–859.

Englund, G. & Olsson, T. I. 1990. Fighting and assessment in the net-spinning caddis larva *Arctopsyche ladogensis*: a test of the sequential assessment game. *Animal Behaviour*, **39**, 55–62.

Enquist, M. & Leimar, O. 1983. Evolution of fighting behavior: decision rules and assessment of relative strength. *Journal of Theoretical Biology*, **102**, 387–410.

Enquist, M., Leimar, O., Ljungberg, T., Mallner, Y. & Segerdahl, N. 1990. A test of the sequential assessment game: fighting in the cichlid fish *Nannacara anomala*. *Animal Behaviour*, **40**, 1–14.

Faber, D. B. & Baylis, J. R. 1993. Effects of body size on agonistic encounters between male jumping spiders (Araneae: Salticidae). *Animal Behaviour*, **45**, 289–299.

Fraser, D. 1974. The behaviour of growing pigs during experimental social encounters. *Journal of Agricultural Science*, **82**, 147–163.

Gammell, M. P. & Hardy, I. C. W. 2003. Contest duration: sizing up the opposition? *Trends in Ecology & Evolution*, **18**, 491–493.

Glass, C. W. & Huntingford, F. A. 1988. Initiation and resolution of fights between swimming crabs (*Liocarcinus depurator*). *Ethology*, **77**, 237–249.

Goubault, M., Batchelor, T. P., Linforth, R. S. T., Taylor, A. J. & Hardy, I. C. W. 2006. Volatile emission by contest losers revealed by real-time chemical analysis. *Proceedings of the Royal Society of London, Series B*, **273**, 2853–2859.

Goubault, M., Batchelor, T. P., Romani, R., Linforth, R. S. T., Fritzsche, M., Francke, W. & Hardy, I. C. W. 2008. Volatile chemical release by bethylid wasps: identity, phylogeny, anatomy and behaviour. *Biological Journal of the Linnean Society*, **94**, 837–852.

Hack, M. A. 1997a. Assessment strategies in the contests of male crickets, *Acheta domesticus* L. *Animal Behaviour*, **53**, 733–747.

Hack, M. A. 1997b. The energetic costs of fighting in the house cricket, *Acheta domesticus* L. *Behavioral Ecology*, **8**, 28–36.

Hammerstein, P. & Parker, G. A. 1982. The asymmetric war of attrition. *Journal of Theoretical Biology*, **96**, 647–682.

Hsu, Y. & Wolf, L. L. 1999. The winner and loser effect: integrating multiple experiences. *Animal Behaviour*, **57**, 903–910.

Hsu, Y., Lee, S. P., Chen, M. H., Yang, S. Y. & Cheng, K. C. 2008. Switching assessment strategy during a contest: fighting in killifish *Kryptolebias marmoratus*. *Animal Behaviour*, **75**, 1641–1649.

Hughes, M. 1996. Size assessment via a visual signal in snapping shrimp. *Behavioral Ecology and Sociobiology*, **38**, 51–57.

Hughes, M. 2000. Deception with honest signals: signal residuals and signal function in snapping shrimp. *Behavioral Ecology*, **11**, 614–623.

Huntingford, F. A. 1984. Some ethical issues raised by studies of predation and aggression. *Animal Behaviour*, **32**, 210–215.

Huyghe, K., Vanhooydonck, B., Scheers, H., Molina-Borja, M. & Van Damme, R. 2005. Morphology, performance and fighting capacity in male lizards, *Gallotia galloti*. *Functional Ecology*, **19**, 800–807.

Hyatt, G. W. & Salmon, M. 1978. Combat in the fiddler crab *Uca pugilator* and *Uca pugnax*: a quantitative analysis. *Behaviour*, **65**, 182–211.

Jennings, D. J., Gammell, M. P., Carlin, C. M. & Hayden, T. J. 2004. Effect of body weight, antler length, resource value and experience on fight duration and intensity in fallow deer. *Animal Behaviour*, **68**, 213–221.

Jennions, M. D. & Backwell, P. R. Y. 1996. Residency and size affect fight duration and outcome in the fiddler crab *Uca annulipes*. *Biological Journal of the Linnean Society*, **57**, 293–306.

Jensen, P. & Yngvesson, J. 1998. Aggression between unacquainted pigs: sequential assessment and effects of familiarity and weight. *Applied Animal Behaviour Science*, **58**, 49–61.

Jenssen, T. A., Orrell, K. S. & Lovern, M. B. 2000. Sexual dimorphisms in aggressive signal structure and use by a polygynous lizard, *Anolis carolinensis*. *Copeia*, **2000**, 140–149.

Johnsson, J. I. & Akerman, A. 1998. Watch and learn: preview of the fighting ability of opponents alters contest behaviour in rainbow trout. *Animal Behaviour*, **56**, 771–776.

Karavanich, C. & Atema, J. 1998. Olfactory recognition of urine signals in dominance fights between male lobster, *Homarus americanus*. *Behaviour*, **135**, 719–730.

Katoh, E., Johnson, M. & Breithaupt, T. 2008. Fighting behaviour and the role of urinary signals in dominance assessment of Norway lobsters, *Nephrops norvegicus*. *Behaviour*, **145**, 1447–1464.

Keeley, E. R. & Grant, J. W. A. 1993. Visual information, resource value, and sequential assessment in convict cichlid (*Cichlasoma nigrofasciatum*) contests. *Behavioral Ecology*, **4**, 345–349.

Kelly, C. D. 2006. Fighting for harems: assessment strategies during male–male contests in the sexually dimorphic Wellington tree weta. *Animal Behaviour*, **72**, 727–736.

Kelly, C. D. & Godin, J. G. J. 2001. Predation risk reduces male–male sexual competition in the Trinidadian guppy (*Poecilia reticulata*). *Behavioral Ecology and Sociobiology*, **51**, 95–100.

Kemp, D. J. 2000. Contest behavior in territorial male butterflies: does size matter? *Behavioral Ecology*, **11**, 591–596.

Kemp, D. J., Alcock, J. & Allen, G. R. 2006. Sequential size assessment and multicomponent decision rules mediate aerial wasp contests. *Animal Behaviour*, **71**, 279–287.

Leimar, O., Austad, S. & Enquist, M. 1991. A test of the sequential assessment game: fighting in the bowl and doily spider *Frontinella pyramitela*. *Evolution*, **45**, 862–874.

Leiser, J. K., Gagliardi, J. L. & Itzkowitz, M. 2004. Does size matter? Assessment and fighting in small and large size-matched pairs of adult male convict cichlids. *Journal of Fish Biology*, **64**, 1339–1350.

Lindstrom, K. 1992. The effect of resource holding potential, nest size and information about resource quality on the outcome of intruder-owner conflicts in the sand goby. *Behavioral Ecology and Sociobiology*, **30**, 53–58.

Manson, J. & Wrangham, R. 1991. Intergroup aggression in chimpanzees and humans. *Current Anthropology*, **32**, 369–390.

Marden, J. H. & Rollins, R. A. 1994. Assessment of energy reserves by damselflies engaged in aerial contests for mating territories. *Animal Behaviour*, **48**, 1023–1030.

Maynard Smith, J. & Parker, G. A. 1976. The logic of asymmetric contests. *Animal Behaviour*, **24**, 159–175.

Mesterton-Gibbons, M., Marden, J. H. & Dugatkin, L. A. 1996. On wars of attrition without assessment. *Journal of Theoretical Biology*, **181**, 65–83.

Moore, J. C., Obbard, D. J., Reuter, C., West, S. A. & Cook, J. M. 2008. Fighting strategies in two species of fig wasp. *Animal Behaviour*, **76**, 315–322.

Morrell, L. J., Backwell, P. R. Y. & Metcalfe, N. B. 2005. Fighting in fiddler crabs *Uca mjoebergi*: what determines duration? *Animal Behaviour*, **70**, 653–662.

Morris, M. R., Gass, L. & Ryan, M. J. 1995. Assessment and individual recognition of opponents in the pygmy swordtails *Xiphophorus nigrensis* and *X. multilineatus*. *Behavioral Ecology and Sociobiology*, **37**, 303–310.

Moya-Larano, J. & Wise, D. H. 2000. Survival regression analysis: a powerful tool for evaluating fighting and assessment. *Animal Behaviour*, **60**, 307–313.

Parker, G. A. 1974. Assessment strategy and evolution of fighting behavior. *Journal of Theoretical Biology*, **47**, 223–243.

Parker, G. A. & Rubenstein, D. I. 1981. Role assessment, reserve strategy, and acquisition of information in asymmetric animal conflicts. *Animal Behaviour*, **29**, 221–240.

Parker, G. A. & Stuart, R. A. 1976. Animal behavior as a strategy optimizer: evolution of resource assessment strategies and optimal emigration thresholds. *American Naturalist*, **110**, 1055–1076.

Pavey, C. R. & Fielder, D. R. 1996. The influence of size differential on agonistic behaviour in the freshwater crayfish, *Cherax cuspidatus* (Decapoda: Parastacidae). *Journal of Zoology*, **238**, 445–457.

Payne, R. J. H. 1998. Gradually escalating fights and displays: the cumulative assessment model. *Animal Behaviour*, **56**, 651–662.

Payne, R. J. H. & Pagel, M. 1996. Escalation and time costs in displays of endurance. *Journal of Theoretical Biology*, **183**, 185–193.

Payne, R. J. H. & Pagel, M. 1997. Why do animals repeat displays? *Animal Behaviour*, **54**, 109–119.

Pomfret, J. C. & Knell, R. J. 2005. Sexual selection and horn allometry in the dung beetle *Euoniticellus intermedius*. *Animal Behaviour*, **71**, 567–576.

Poole, J. H. 1989. Announcing intent: the aggressive state of musth in African elephants. *Animal Behaviour*, **37**, 140–152.

Pratt, A. E., McLain, D. K. & Lathrop, G. R. 2003. The assessment game in sand fiddler crab contests for breeding burrows. *Animal Behaviour*, **65**, 945–955.

Prenter, J., Elwood, R. W. & Taylor, P. W. 2006. Self-assessment by males during energetically costly contests over precopula females in amphipods. *Animal Behaviour*, **72**, 861–868.

Prenter, J., Taylor, P. W. & Elwood, R. W. 2008. Large body size for winning and large swords for winning quickly in swordtail males, *Xiphophorus helleri*. *Animal Behaviour*, **75**, 1981–1987.

Renison, D., Boersma, D. & Martella, M. B. 2002. Winning and losing: causes for variability in outcome of fights in male Magellanic penguins (*Spheniscus magellanicus*). *Behavioral Ecology*, **13**, 462–466.

Riechert, S. E. 1978. Games spiders play: behavioral variability in territorial disputes. *Behavioral Ecology and Sociobiology*, **3**, 135–162.

Rillich, J., Schildberger, K. & Stevenson, P. A. 2007. Assessment strategy of fighting crickets revealed by manipulating information exchange. *Animal Behaviour*, **74**, 823–836.

Rosenberg, R. H. & Enquist, M. 1991. Contest behaviour in Weidemeyer's admiral butterfly *Limenitis weidemeyerii* (Nymphalidae): the effect of size and residency. *Animal Behaviour*, **42**, 805–811.

Rushen, J. 1987. A difference in weight reduces fighting when unacquainted newly weaned pigs first meet. *Canadian Journal of Animal Science*, **67**, 951–960.

Rushen, J. 1988. Assessment of fighting ability or simple habituation: what causes young pigs (*Sus scrofa*) to stop fighting? *Aggressive Behavior*, **14**, 155–167.

Sell, A., Cosmides, L., Tooby, J., Sznycer, D., Von Rueden, C. & Gurven, M. 2009. Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society of London, Series B*, **276**, 575–584.

Smallegange, I. M., Sabelis, M. W. & Van Der Meer, J. 2007. Assessment games in shore crab fights. *Journal of Experimental Marine Biology and Ecology*, **351**, 255–266.

Smith, I. P., Huntingford, F. A., Atkinson, R. J. & Taylor, A. C. 1994. Strategic decisions during agonistic behaviour in the velvet swimming crab, *Necora puber* (L.). *Animal Behaviour*, **47**, 885–894.

Sneddon, L. U., Huntingford, F. A. & Taylor, A. C. 1997. Weapon size versus body size as a predictor of winning in fights between shore crabs, *Carcinus maenas* (L.). *Behavioral Ecology and Sociobiology*, **41**, 237–242.

Stamps, J. A. & Krishnan, V. V. 1994. Territory acquisition in lizards: 1. First encounters. *Animal Behaviour*, **47**, 1375–1385.

Stokkebo, S. & Hardy, I. C. W. 2000. The importance of being gravid: egg load and contest outcome in a parasitoid wasp. *Animal Behaviour*, **59**, 1111–1118.

Stuart-Fox, D. 2006. Testing game theory models: fighting ability and decision rules in chameleon contests. *Proceedings of the Royal Society of London, Series B*, **273**, 1555–1561.

Stuart-Fox, D. M. & Johnston, G. R. 2005. Experience overrides colour in lizard contests. *Behaviour*, **142**, 329–350.

Stuart-Fox, D. M. & Whiting, M. J. 2005. Male dwarf chameleons assess risk of courting large, aggressive females. *Biology Letters*, **1**, 231–234.

Stuart-Fox, D. M., Firth, D., Moussalli, A. & Whiting, M. J. 2006. Multiple signals in chameleon contests: designing and analysing animal contests as a tournament. *Animal Behaviour*, **71**, 1263–1271.

Stutt, A. D. & Wilmer, P. 1998. Territorial defence in speckled wood butterflies: do the hottest males always win? *Animal Behaviour*, **55**, 1341–1347.

Taylor, P. W. & Elwood, R. W. 2003. The mismeasure of animal contests. *Animal Behaviour*, **65**, 1195–1202.

Taylor, P. W., Hasson, O. & Clark, D. L. 2001. Initiation and resolution of jumping spider contests: roles for size, proximity and early detection of rivals. *Behavioral Ecology and Sociobiology*, **50**, 403–413.

Tibbetts, E. A. & Curtis, T. R. 2007. Rearing conditions influence quality signals but not individual identity signals in *Polistes* wasps. *Behavioral Ecology*, **18**, 602–607.

Tibbetts, E. A. & Dale, J. 2004. A socially enforced signal of quality in a paper wasp. *Nature*, **432**, 218–222.

Tibbetts, E. A. & Lindsay, R. 2008. Visual signals of status and rival assessment in *Polistes dominulus* paper wasps. *Biology Letters*, **4**, 237–239.

Triefenbach, F. A. & Zakon, H. H. 2008. Changes in signalling during agonistic interactions between male weakly electric knifefish, *Apteronotus leptorhynchus*. *Animal Behaviour*, **75**, 1263–1272.

Vanhooydonck, B., Herrel, A., Van Damme, R., Meyers, J. J. & Irschick, D. J. 2005. The relationship between dewlap size and performance changes with age and sex in a green anole (*Anolis carolinensis*) lizard population. *Behavioral Ecology and Sociobiology*, **59**, 157–165.

Verrell, P. A. 1986. Wrestling in the red-spotted newt (*Notophthalmus viridescens*): resource value and contestant asymmetry determine contest duration and outcome. *Animal Behaviour*, **34**, 398–402.

Ward, P. I. 1983. Advantages and a disadvantage of large size for male *Gammarus pulex* (Crustacea, Amphipoda). *Behavioral Ecology and Sociobiology*, **14**, 69–76.

Wells, M. S. 1988. Effects of body size and resource value on fighting behaviour in a jumping spider. *Animal Behaviour*, **36**, 321–326.

Whitehouse, M. E. A. 1997. Experience influences male–male contests in the spider *Argyrodes antipodiana* (Theridiidae: Araneae). *Animal Behaviour*, **53**, 913–923.

Wilson, J. D. 1992. Correlates of agonistic display by great tits *Parus major*. *Behaviour*, **121**, 168–214.

Wise, S. E. & jaeger, R. G. 1998. The influence of tail autotomy on agonistic behaviour in a territorial salamander. *Animal Behaviour*, **55**, 1707–1716.

Zamudio, K. R., Huey, R. B. & Crill, W. D. 1995. Bigger isn't always better: body-size, developmental and parental temperature and male territorial success in *Drosophila melanogaster*. *Animal Behaviour*, **49**, 671–677.

Zucker, N. & Murray, L. 1996. Determinants of dominance in the tree lizard *Urosaurus ornatus*: the relative importance of mass, previous experience and coloration. *Ethology*, **102**, 812–825.

Zulandt-Schneider, R. A., Huber, R. & Moore, P. A. 2001. Individual and status recognition in the crayfish, *Orconectes rusticus*: the effects of urine release on fight dynamics. *Behaviour*, **138**, 137–153.