Title
Coalitions in theory and reality: a review of pertinent variables and processes

Permalink
https://escholarship.org/uc/item/6zf593zz

Journal
Behaviour, 152(1)

ISSN
0005-7959

Authors
Bissonnette, Annie
Perry, Susan
Barrett, Louise
et al.

Publication Date
2015

DOI
10.1163/1568539x-00003241

Peer reviewed
Coalitions in theory and reality: A review of pertinent variables and processes

Annie Bissonnette¹,², Susan Perry³, Louise Barrett⁴,⁵, John C. Mitani⁶, Mark Flinn⁷, Sergey Gavrilets⁸ & Frans B. M. de Waal⁹

¹Courant Research Centre Evolution of Social Behaviour, Georg-August-University Göttingen, Kellnerweg 6, 37077 Göttingen, Germany
²Anthropological Institute and Museum, University of Zürich, 190 Winterthurerstrasse, 8057 Zürich, Switzerland
³Department of Anthropology, Behavior, Evolution and Culture Program, University of California-Los Angeles, 375 Portola Plaza, Los Angeles, CA 90095-1553, USA
⁴Department of Psychology, University of Lethbridge, 4401, University Drive West, Alberta T1K 3M4, Canada
⁵Department of Anthropology, University of Michigan, 1085 South University Avenue, Ann Arbor, MI 48109-1107, USA
⁶Department of Anthropology, University of Missouri, 107 Swallow, Columbia, MO 65211, USA
⁷Department of Ecology and Evolutionary Biology and Department of Mathematics, National Institute for Mathematical and Biological Synthesis, University of Tennessee, Knoxville, TN 37996, USA
⁸Living Links, Yerkes National Primate Research Center, Emory University, 954 North Gatewood Rd., Atlanta, GA 30329, USA

Short title: Coalitions in theory and reality

Email corresponding author: a.bissonnette@aim.uzh.ch
Abstract

Coalitions and alliances are ubiquitous in humans and many other mammals, being part of the fabric of complex social systems. Field biologists and ethologists have accumulated a vast amount of data on coalition and alliance formation, while theoretical biologists have developed modeling approaches. With the accumulation of empirical data and sophisticated theory, we are now potentially able to answer a host of questions about how coalitions emerge and are maintained in a population over time, and how the psychology of this type of cooperation evolved. Progress can only be achieved, however, by effectively bridging the communication gap that currently exists between empiricists and theoreticians. In this paper, we aim to do so by asking three questions: (1) What are the primary questions addressed by theoreticians interested in coalition formation, and what are the main building blocks of their models? (2) Do empirical observations support the assumptions of current models, and if not, how can we improve this situation? (3) Has theoretical work led to a better understanding of coalition formation, and what are the most profitable lines of inquiry for the future? Our overarching goal is to promote the integration of theoretical and field biology by motivating empirical scientists to collect data on aspects of coalition formation that are currently poorly quantified and to encourage theoreticians to develop a comprehensive theory of coalition formation that is testable under real-world conditions.

Keywords: Alliances, mathematical models, empirical data, primates, recruitment, fitness, dominance, cooperation
Introduction

Ideally, theory and empirical data should build on each other in a series of mutually reinforcing cycles, generating ever more powerful predictive models that could be tested empirically under a wide range of laboratory and field conditions. The greater precision of our models, and the more focused data sets that would then accumulate, would enable us to identify the key elements underlying a particular behavioral phenomenon. This ideal is rarely met, however, and theoreticians and empirical scientists all too often talk past one another. Part of the reason is a difference in the generality of issues under study: many animal behavior researchers stay close to the empirical material they gather, often focusing on just one or two groups of animals for practical reasons, while asking quite specific questions regarding the observed patterns. In contrast, theoreticians often use formal description to arrive at more general underlying principles of a given phenomenon that can be broadly applied across a variety of taxa and social systems. They do not generally formulate their models in a way that facilitates empirical investigation via precise operationalization.

Here we consider the case of coalition formation, an active area of research that seems ready for mathematical modeling. Naturalistic observations are crucial for documenting its distribution across taxa, the variety of forms it takes, the function it serves, and the degree of cognitive complexity it requires. As we will discuss, however, the complexity of the patterns and processes revealed by these naturalistic studies demonstrates the need for a formal theoretical framework to fully understand the evolution and maintenance of coalition formation across the animal kingdom. There is currently a whole suite of modeling approaches available to address such fundamental questions as: Why do some species commonly form coalitions whereas others do not? What are the potential fitness consequences of not participating in all possible coalition opportunities? Can apparently "cognitively complex" coalitional behavior emerge from simple behavioral rules? Unfortunately, as noted above, the exchange of ideas between empiricists and theoreticians is often hindered by their different goals, background, and use of terminology (Mesterton-Gibbons et al., 2010). Here, we
attempt to forge an alliance between theoretical and empirical approaches as the first step toward
the long-term goal of achieving a more comprehensive theory of coalition formation.

We start by briefly describing patterns of coalition formation within and across species. We next
summarize the primary questions addressed by theoreticians, and identify the main building blocks
of their models. Then we selectively review the literature on coalition formation with the explicit aim
of contrasting empirical data with the assumptions made by theoretical models, pinpointing where
these might be overly simplistic, and identifying important parameters that have been poorly
quantified in the field. As the literature in this area remains heavily skewed toward the primates
(Smith et al., 2010), the empirical data presented here are derived mainly from this taxon. Finally, we
propose some profitable lines of inquiry for future theoretical and empirical research on coalition
formation.

A Brief Primer on Coalition Formation

The term “coalition” is used in various ways (e.g. an alliance between political parties, Mann, 1993)
but, in ethology and behavioral biology, the term typically refers to two (or more) individuals acting
together against a third party in an aggressive or competitive context (Harcourt & de Waal, 1992;
Figure 1). The term "alliance," on the other hand, generally refers to a long term relationship
characterized by repeated coalitions and high levels of association between the partners, which
means that when one partner is in trouble, his or her ally is often nearby and able to help (“alliance
behavior based on coalitions”: cf. Feh, 1999; de Waal & Harcourt, 1992). Human coalitions and
alliances range from cooperation within a single family to acts of support between nation-states
(Chagnon, 1988b; Bowles, 2009; Apicella et al., 2012). On a smaller scale, coalitions and alliances are
common in our closest relatives, including for example the chimpanzee (Pan troglodytes, review in
Muller & Mitani, 2005), as well as a number of non-primate species (reviewed by Harcourt & de
Waal, 1992; Smith et al., 2010).
Animal coalitions and alliances are variable in duration and stability. Coalitions can occur opportunistically on a case-by-case basis (e.g. Bissonnette, 2009; Smith et al., 2010), or they can be observed frequently among the same partners as part of an enduring, long-term relationship (e.g. de Waal, 1982; Packer & Pusey, 1982; Smuts, 1985; Goodall, 1986; Noë, 1986b; Caro & Collins, 1987; Connor et al., 1992; Feh, 1999; Boesch & Boesch-Achermann, 2000; Connor et al., 2001; Boinski et al., 2005; Schülke et al., 2010). One individual may establish year-long, stable alliances with one or a few partners and also participate in opportunistic coalitions with other group members in which loyalties change regularly (Riss & Goodall, 1977; de Waal, 1982; Goodall, 1986; Noë, 1990; Uehara et al., 1994; Connor et al., 2000; Sijtsema et al., 2009). For example, Nishida (1983) and de Waal (1984) have argued that male chimpanzees intervene opportunistically in "political" conflicts and may turn against former allies when it is beneficial to do so.

In group-living species, coalitions and alliances are important in both intergroup and intragroup competition. Coalitions for competition against conspecifics from other groups (hereafter “group-wide coalitions”) occur when the opponents belong to different social groups (or family units, e.g. Scheiber et al., 2005), whereas within-group coalitions occur when two or more members of the same group compete against at least one other “in-group” member. Group-wide coalitions can be offensive if individuals join forces to oust incumbent male(s) and take over a group (e.g. Pope, 1990; Rood, 1990; Packer et al., 1991; Waser et al., 1994; Ostner & Kappeler, 2004; Wilson et al., 2014), or to conduct a lethal attack on members of a neighbouring group (reviewed by Wrangham, 1999). They can also be defensive, as when group members form a coalition to repel single challengers and/or an out-group coalition (e.g. Feh, 1999). Group-wide coalitions can be very large (e.g. average size in chimpanzees: 13, Watts & Mitani, 2001; hyenas: 14, Smith et al., 2010; human band and tribal level raiding parties range from 10 - 100, Walker & Bailey, 2013; Mathew & Boyd, 2011), in contrast to within-group coalitions, which often involve only two or three partners (reviewed by Smith et al., 2010). It seems likely that group-wide coalition formation provides the only context in which the
opposing parties are not mutually dependent on each other (except in dolphins and especially humans, where interdependence extends beyond the level of the group, e.g. Connor et al., 2001; Walker et al., 2011): they usually occur in a clear and unambiguous competitive context and thus are predictably hostile, in contrast to the interactions that occur within groups (Connor et al., 1992; Harcourt, 1992; see below). Group-wide and within-group coalitions, however, are not mutually exclusive phenomena: for example, the need for in-group members to cooperate during between-group encounters is known to profoundly affect the social dynamics of groups, and hence the formation of within-group coalitions (humans: Alexander, 1990; 2014; Flinn et al., 2005; Choi & Bowles, 2007; Bowles, 2009; Rusch 2014; other animals: Watts & Mitani, 2001). Group-wide coalitions have also been suggested as one possible evolutionary precursor of within-group coalitions in animals (Harcourt, 1992), although a formal test of this hypothesis is lacking. Despite the likely interdependence between intra- and inter-group coalition formation, the emphasis of this review is on within-group coalitions, as these have received much more attention both theoretically and empirically.

Research on within-group coalitions has historically focused on primates (Kawai, 1958; Kawamura, 1958; Hall & DeVore, 1965; Kummer, 1967; Packer, 1977; Riss & Goodall, 1977; de Waal, 1978b). More recent empirical work, however, has revealed that a wide range of mammalian species and some birds engage in coalitionary behaviour (reviewed by Smith et al., 2010), including several social carnivores (e.g. hyenas, Crocuta crocuta: Zabel et al., 1992; Engh et al., 2002; Wahaj et al., 2004; ring-tailed coatis, Nasua nasua: Romero & Aureli, 2008; African wild dog, Lycaon pictus: de Villiers et al., 2003; Domestic dogs, Canis lupus familiaris: Ward et al., 2009), some artiodactyls (e.g. fallow deer, Dama dama: Jennings et al., 2009; feral sheep, Ovis aries: Rowell & Rowell, 1993), some perissodactyls (e.g. plain zebra, Equus quagga: Schilder, 1990), one proboscid (African elephant, Loxodonta africana: Lee, 1987), and some birds (e.g. ravens, Corvus frugilegus: Seed et al., 2007; Corvus corax: Fraser & Bugnyar, 2012).
The pattern of within-group coalition formation most commonly reported in the literature is that of one individual intervening in an on-going conflict and aligning itself with one of the original opponents (variously referred as “agonistic intervention”, “agonistic aid”, “fight interference”, “support choice”, de Waal & Harcourt, 1992). Individuals also sometimes join forces to simultaneously attack the same target prior to the initiation of a conflict with that target (i.e. instantaneous or parallel coalitions, cf. Noë, 1986c; e.g. de Waal, 1982; Bissonnette, 2009). For example, most chimpanzee male coalitions at Ngogo formed spontaneously or when the males jointly retaliated after another male displayed in the vicinity (Watts, 2002). The term coalition covers both of these interaction patterns, but excludes cases where an intervening animal breaks up an agonistic interaction between two others without siding or “acting together” with either putative combatant (referred to as “pacifying”, “neutral”, “policing”, “peaceful” or “impartial” interventions, e.g. Bernstein, 1964; Boehm, 1994; Petit & Thierry, 1994; Roeder et al., 2002; Flack et al., 2006; Rudolf von Rohr et al., 2012).

Three configurations of within-group coalitions formed by two partners against a single target have been recognized (cf. Chapais, 1995; van Schaik et al., 2006; see Chapais 1995 for various examples in primates):

(i) “all-down” (or “conservative”) when the coalition members outrank their target;

(ii) “bridging” when one partner ranks above and one partner ranks below the target;

(iii) “all-up” (or “revolutionary”) when both partners rank below their target.

Although some variation within- and across species has been reported (e.g. Harcourt, 1992; Boesch & Boesch-Achermann, 2000; Young et al., 2014), all-down coalitions appear to be the most common coalitionary pattern observed in animals when the beneficiary is not related to the supporter (e.g. Harcourt & Stewart, 1989; Silk, 1993; Smith et al., 2010).

As we will see below, coalitions can serve different functions (van Schaik et al., 2006); they can be observed among different sex and age classes (reviewed in Chapais, 1995); they can involve kin (e.g.
Riss & Goodall, 1977; Chagnon & Bugos, 1979b; Wahaj et al., 2004), non-kin (e.g. Vigilant et al., 2001; Langergraber et al., 2007; Schuelke et al., 2010), and friends or acquaintances (Hruschka & Henrich, 2006; Hruschka, 2010). As this brief review makes clear, coalitionary patterns are varied, and understanding this diversity would be greatly enhanced by formal modeling that would allow us to identify the general conditions under which coalition formation is expected to evolve, characterize the degree of variability expected, and predict when coalitions should, and should not, occur. Below, we summarize existing modeling approaches and highlight their main findings and predictions to date.

**Modeling coalitions and alliances**

**General theoretical perspectives**

There exist four major mathematical theories that can provide insight into how coalitions and alliances work (for more technical details see a recent review by Mesterton-Gibbons et al., 2011). Of these, two have been most commonly applied to coalition formation. First, the *theory of endogenous coalition formation* has been developed in the economics literature (Tullock, 1980; Skaperdas, 1996; Skaperdas, 1998; Noh, 2002; Esteban & Sákovics, 2003; Garfinkel, 2004; Bloch et al., 2006; Sánchez-Pagés, 2007; Konrad & Kovenock, 2009; Stamatopoulos et al., 2009; Tan & Wang, 2010). This theory aims to predict coalitionary strategies that maximize individual fitness given a particular (ecological) situation, the costs and benefits of different actions, and the information available to individuals.

Most existing models of coalition formation in animals belong to this general framework, which overlaps with evolutionary game theory (Noë, 1990; 1992; 1994; Pandit & van Schaik, 2003; van Schaik et al., 2004; 2006; Connor & Whitehead, 2005; Whitehead & Connor, 2005; Mesterton-Gibbons & Sherratt, 2007; Broom et al., 2009; Okasha, 2009). We will discuss a number of these models in more detail below.
Second, the *theory of cooperative games* models situations where players that have pooled their resources attempt to come to an agreement on how to divide the spoils (von Neumann & Morgenstern, 1947; Kahan & Rapoport, 1984; Myerson, 1991; Ray, 2007; Gilles, 2010). Coalitionary theory using this perspective on cooperation predicts which coalitions can form and how the benefits are distributed. Cooperative game theory proved to be useful in the study of human behavior, but has rarely been applied to coalitions in non-human animals. This is at least partly because it remains unclear how animals could negotiate without language and how any agreement could be enforced (McNamara et al., 2006; Akcay & Roughgarden, 2007; Binmore, 2010). One notable exception is Noë (1990), who used a veto game from cooperative game theory to explain patterns of coalition formation among adult males in a group of wild baboons. In the veto game, there is a single “strong” player who is able to claim (i.e., without negotiating) most of the spoils by playing other individuals against each other. The power of a veto player may also be based on a unique skill, or possession of a unique resource. A key player playing off two partners against each other is a situation that has been described for both captive and wild chimpanzees (de Waal, 1982; Nishida, 1983; Muller & Mitani, 2005), and primate negotiation without language is discussed by de Waal (1996) and Melis et al. (2009).

Third, the *theory of dynamic linking and network formation* focuses on dyadic or polyadic games played by individuals in which the outcomes of previous interactions determine who interacts with whom in the future (Skyrms & Pemantle, 2000; Pemantle & Skyrms, 2004a; 2004b; Hruschka & Henrich, 2006; Santos et al., 2006; Pacheco et al., 2006; 2008; Gavrilets et al., 2008; Aktipis, 2011). We discuss one application of this theory below.

Fourth, *coagulation-fragmentation theory* studies the dynamics of the population structure given some simple rules concerning the mechanisms by which individuals join and leave groups (which in our context can be thought of as representing coalitions). Its origins go back to a 1917 study of polymers by Smoluchowski (Wattis, 2006), and have only recently found a way into the social sciences (Eguíluz & Zimmermann, 2000; Bohorquez et al., 2009; but see: Cohen, 1972). This
promising approach will not be discussed further in this paper because it has not yet been applied to coalitionary behavior in animals.

Specific goals of theoretical models

From a biological perspective, there are three major sets of questions that modelers aim to answer. First, which strategies maximize individual fitness? Second, which coalitionary patterns are expected to emerge, and how do they affect individual fitness and behavior? Third, how do the behavioral rules used by individuals evolve and what are the individual and group level consequences of such evolution? To illustrate how theoreticians address these questions, we discuss several recent models in more detail and present their central assumptions in Table 1.

Fitness maximization

An overwhelming majority of coalitionary models focuses on determining which of a possible set of coalitionary strategies maximizes individual fitness, given specific assumptions about the costs and benefits to individuals, as well as the group composition and the information available to individuals. Typically these models involve only three individuals. One of the oldest models is the one of "minimally winning coalitions" of Riker (1962), i.e. rational players form coalitions in which the pooled resources are as small as possible yet sufficient to win. In a more detailed analysis, Stamatopoulos et al. (2009) investigated how differences in strength between three individuals competing for a unit of divisible resource affect whether coalitions of two against one will emerge and which individuals will form them. In their model, they assumed that: (a) the strength of a coalition is given by the sum of the strengths of its participants, (b) the probability of winning a fight (and thus obtaining the resource) is given by a logistic function of the difference in strengths, (c) a winning coalition divides the reward in proportion to the strengths of the individuals in the coalition, and (d) the costs of fighting are absent. Using additional assumptions about how the reward is divided if no coalition is formed, Stamatopoulos et al. (2009) showed that if the strongest individual is sufficiently strong relative to the weakest members, the latter should form an all-up coalition
against the former. If the strengths of the two strongest members are relatively close to that of the weakest member, they should form an all-down coalition. In all other cases, the strongest and the weakest players should form a bridging coalition against the medium-strength player. Consistent with one of the model's prediction, Benenson et al. (2009) found that human subjects are increasingly willing to form coalitions as their own level of perceived strength diminishes.

Broom et al. (2009) studied the effects of coalitionary support by kin on the stability of dominance hierarchies in group-living animals. Their model considers a conflict between a high-ranking individual and a challenger seeking to replace it in the hierarchy. The conflict is observed by a third individual who can decide to help one of the two parties. Helpers derive only indirect fitness benefits as a consequence of their relatedness to the main actors. Specifying the costs and benefits of different actions, Broom et al.’s (2009) model predicts who should attack whom under which circumstances. An interesting prediction is that increased variance in relatedness within a group, but not the mean relatedness, makes formation of coalitions more plausible (i.e. if variance in relatedness is low, regardless of the mean degree of relatedness among all individuals, coalitions will not occur).

Using a similar approach, van Schaik and colleagues (Pandit & van Schaik, 2003; van Schaik et al., 2004; 2006) aimed to predict the type of coalitions formed (i.e. all-up, bridging, or all-down; rank-changing or leveling) in a group of males competing for fertilization opportunities. They considered primate males living in groups of an arbitrary size, where individuals ranked in a linear hierarchy compete over increases in ranks. van Schaik et al. (2004; 2006) assumed: (a) that the relationship between the rank and fitness pay-off is exponential, (b) that coalitions are formed only against a single target, (c) that coalition strength is given by a sum of the strengths of its participants, and (d) that the stronger party always wins. Using an idealized cost parameter, van Schaik et al. (2004; 2006) identified regions of parameter space where coalitions were both feasible (i.e., are stronger than the target) and profitable. Building on these results, van Schaik et al. (2004; 2006) made a number of testable predictions. For example, they predicted that all-up rank-changing coalitions should (a) arise
only in groups with a medium degree of despotism, (b) be a small coalition (two or three members),
(c) target the top individual, and (d) involve individuals ranking just below the target. Recent
empirical tests have provided support for some predictions of the Pandit/van Schaik model in
primate males (Jones, 2005; Bissonnette, 2009; Higham & Maestripieri, 2010; but see Kulik et al.,
2012), although important limitations of the approach have also been discussed (Bissonnette, 2009;
Berghänel et al., 2010; Young et al., 2014).

There is a general criticism of the optimization approach embodied in fitness maximization
models. Specifically, the applicability of this approach hinges on the assumption that the optimal
strategies identified in the mathematical models are (or closely approximate) the strategies used by
real individuals. How can animals know what the best strategies are? One common answer is that
natural selection is able to find the best strategy so that organisms have evolved to develop an
optimum “coalitionary psychology”. Given the multiplicity and interdependence of possible
behaviors and the complexity of resulting evolutionary dynamics in a coalitionary context, this
answer is not entirely satisfactory (McNamara & Houston, 2009). Also, it is possible that the
optimization problem has multiple alternative solutions, which means that the most relevant
solution is not always obvious.

From behavioral rules to coalitionary patterns

An alternative approach is to predict group level coalitionary patterns and their effects on individual
fitness and behavior, given the assumptions listed above, as well as some additional assumptions
about how individuals interact in dyadic and polyadic conflicts. For example, using the theory of
dynamic linking, Gavrilets et al. (2008) examine the theoretical plausibility of an argument about the
role of coalitions in the “egalitarian revolution” (Boehm, 1999), which is suggested to have resulted
in a transition from the hierarchical societies characteristic of chimpanzees to the egalitarian
societies seen among human hunter-gatherers. Gavrilets et al. (2008) found support for Boehm’s
arguments and further suggested that the transition required a change in cognitive abilities. Gavrilets
et al. (2008) assumed that decisions regarding whether to join a coalition were based on a simple heuristic rule that utilized the relative “affinities” of a potential helper to the two individuals engaged in conflict. These “affinities” reflected the history of past interactions and changed via a process akin to reinforcement learning: that is, they increased following a mutually beneficial interaction and decreased following an agonistic interaction. Affinities also continuously decayed to zero reflecting a loss of memory by the helpers and actors (or alternatively, the discounting of distant events in the past). Affinities thus represent a simple alternative to explicit fitness considerations. A particularly attractive feature of Gavrilets et al.’s (2008) model and similar approaches (e.g. Pacheco et al., 2008) is that they are scalable, i.e. they can easily be generalized to larger groups and multiple coalitions. Of course this would require individuals to be able to memorize a larger number of affinities. One serious limitation to this approach must also be emphasized, however: there is no guarantee that the heuristic rules used in this and similar models are evolutionarily stable. Consequently, the “coalitionary psychology” postulated in these models is, potentially, biologically irrelevant. This is clearly an area on which empirical research can shed more light by testing which decision rules are used in coalitionary conflicts.

Evolution of behavioral rules and its consequences

Analysis of evolutionarily stable strategies has been an extremely important tool in theoretical studies of animal behavior (Maynard Smith, 1982). There are, however, only a few models that explicitly study the emergence of evolutionarily stable behavioral rules in a coalitionary context (thus combining the two approaches outlined above). These models make specific assumptions with respect to possible coalitionary strategies, their costs and benefits, the information available to individuals, and the behavioral rules followed by individuals. Mesterton-Gibbons and Sherratt (2007), for example, consider competition among three individuals in a setup similar to that of Stamatopoulos et al. (2009) and Broom et al. (2009). A crucial novel feature of their model, however, is the assumption that individuals differ with respect to a genetically controlled trait (“alliance threshold”) that determines whether an individual seeks coalitionary aid (if his strength falls below
threshold) or goes it alone (if his strength exceeds threshold). Mesterton-Gibbons and Sherratt (2007) also allow for incomplete information about the strengths of other individuals, for dependence of fighting costs on the strengths of competitors, and for synergy/antergy between individual strengths in determining pooled strength. They investigated how evolutionarily stable alliance threshold values (which determine evolved “coalitionary psychology”) depend on various parameters of their model. In particular, they predicted that 2-versus-1 coalitions are more likely to occur when (a) the reliability of strength as a predictor of contest outcome is high, (b) there is high variation in fighting strengths, and counterintuitively (c) when the effective strength of a coalition is less than the sum of its individual strengths (see also Mesterton-Gibbons & Sherratt, 2009; Gavrilets, 2012). Using similar methods for identifying evolutionary stable behaviors in various coalitionary contexts would greatly inform the theory (but will likely be quite challenging mathematically).