

## **Evolutionary psychologists need to distinguish between the evolutionary process, ancestral selection pressures, and psychological mechanisms**

**John Tooby and Leda Cosmides**

*Department of Psychology, Stanford University, Stanford, CA 94305*  
*Electronic mail: leda@psych.stanford.edu*

Although many of the experiments reviewed by Caporael et al. are interesting, and we are sympathetic to some of the intuitions that motivate the target article (e.g., that "selfishness" and "self-interest" as phrases can sometimes be misleading as characterizations of human psychological mechanisms), the authors' misconstrual of much of modern evolutionary theory makes their interpretation of their own work problematic. In particular, their categorization of modern evolutionary theory as an "egoistic incentive" (EI) theory misconstrues claims about the evolutionary process itself as claims about the psychological mechanisms that are the shaped product of the evolutionary process. Because Caporael et al. are not alone in this confusion, but rather hold misconceptions that persist both within and outside of the evolutionary community, it is worth dwelling on exactly where arguments about whether humans are "basically selfish" or (the authors' alternative) "basically social" go wrong.

In approaching a given species' behavior from an adaptationist perspective, evolutionary analysis requires three nested but distinct levels. These are:

(1) Models of the evolutionary process (involving definitions of fitness, selection, adaptation, genes, and the role of stochastic factors, and general models of such topics as kin-directed altruism, reciprocation, sexual recombination, and sexual selection);

(2) An analysis of how these principles were manifested as a species-specific array of selection pressures, refracted through the specific ecological, social, genetic, phylogenetic, and informational circumstances experienced along a given species' evolutionary lineage;

(3) a description of the species' innate adaptive specializations that evolved to solve the problems posed by the species-specific array of ancestral selection pressures.

The advances over the last three decades in the sophistication and power of evolutionary theory have justifiably generated an enormous amount of excitement, and although much still remains to be done, this first level is the best developed of the three (Dawkins 1976; 1982; Hamilton 1964; Maynard Smith 1982a; Williams 1966). The second level, the reconstruction of the array of selection pressures encountered over hominid evolution (in the human case), has made a limited amount of progress (see discussion in Tooby & DeVore 1986). The third level, the mapping of human psychological mechanisms as adaptations, is in its infancy and depends strongly on the other two for its success. The three levels fit together in a tightly structured way, and all are necessary to address any behavioral issue with confidence. The relative maturity of evolutionary theory compared to the other two levels has led the overeager to try to leap directly from evolutionary theory to psychology, in the belief that evolutionary theory by itself constituted a theory of psychology (Cosmides & Tooby 1987).

To understand exactly why one cannot leap from evolutionary theory to psychological theory, consider a thought experiment involving members of a hypothetical species of fig wasp whose life history leads them to enter, and spend all of their adult lives, encased in a single fig. Further assume that after recruiting a single mate during their juvenile phase, each pair colonizes a different fig. Adults would always interact only with their mates, and would never encounter another conspecific. The evolved mechanisms regulating altruism towards the mate would act as if they "valued" the welfare of the mate as highly as the individual they were in, because the death or disability of the mate would be reproductive death for the wasp (see discussion in Alexander 1987 on marriage). Moreover, such mechanisms would not evolve to dole out assistance differentially depending on the recipient, but would be indiscriminately altruistic, because the only conspecific ever encountered by an adult was its mate, and so indiscriminate aid was never an error. Finally, imagine researchers who brought such wasps into the lab to investigate whether wasps were "selfish" inclusive fitness maximizers or whether the "selfish gene" theory was bankrupt. In raising such wasps with others, and running experiments on them in various social situations, they would discover that the wasps were, in contradiction to "theory," perfectly altruistic towards non-relatives (thus ruling out inclusive fitness explanations of altruism) and toward "strangers" (ruling out reciprocity or mate cooperation as an explanation of altruism). Thus, a little hard-headed empiricism would dispel all these "selfish gene" theories. Similarly, Caporaël et al. experiment on humans by putting them in situations where, for example, the "subjects were strangers," "their choices were anonymous," "they made a single decision," and "interaction among group members was prevented before and after each session" – i.e., in evolutionarily unprecedented situations, completely unlike anything humans would have regularly encountered during their Pleistocene evolution, and therefore situations to which their psychological mechanisms are not adapted.

Clearly, without a theory of the selection pressures operating in ancestral environments, evolutionary theory has little to say about psychology, and equally, psychology without this necessary intermediate cannot serve as a test for "selfish gene" approaches (as Caporaël et al. appear to believe). This intervening level is necessary to make models of psychological mechanisms relate to evolutionary theory and vice versa. When these three levels are not kept clearly in mind, and models are not evaluated in terms of the appropriate level of analysis, confusion abounds. What Caporaël et al. call "selfish gene" theories are the best existing characterizations of how the *process* of natural selection operates: to refute these theories, the authors would have to propose some new theory of the process of natural selection, or some overlooked element in present models of fitness, and this is exactly what the authors do not do. Instead, they misinterpret such theories as theories about motivation –

something like: Under all circumstances, evolutionarily unprecedented or not, a human cognitively represents inclusive fitness, and "selfishly" has as a goal the desire to maximize it. In all fairness, one must admit that some evolutionary researchers treat evolutionary theory this way as well (i.e., they treat humans as fitness strivers), although more often this is simply a metaphorical conceptual shorthand, used to avoid cumbersome constructions (as Dawkins [1976; 1982] takes great pains to make clear). The requirement to find "individualistic payoff," that is, enhanced genetic propagation, as the designed consequence for any psychological adaptation is not a level 3 statement about psychological goals. Rather, it is the expression of the central requirement of the theory of natural selection (level 1): adaptations exist, and have the form they do, because they enhanced genetic propagation (fitness) in ancestral conditions. No matter what psychological mechanism is proposed, it is incumbent on the researcher to show how genes coding for it could have spread through the population. Discovering how they accomplished this is not a luxury (let alone a metatheoretical prejudice deriving from Western cultural beliefs), but is rather an inescapable logical necessity deriving from the structure of evolution by natural selection. Caporaël et al.'s "sociality hypothesis" is a vaguely expressed theory of level 2 (species-specific selection pressures), introduced to account for the in-group bias phenomena they and others have encountered (a psychological phenomenon, level 3). They incorrectly see their level 2 hypothesis as exempting them somehow from showing how in-group bias (a level 3 phenomenon) leads to "individualistic payoffs," that is, enhanced fitness (the level 1 requirement) in ancestral environments (the level 2 context). The logically required attempt by others to find level 1 and level 2 explanations for group living (i.e., its fitness payoffs) is interpreted by the authors to mean that other researchers are somehow ignorant of the fact that humans evolved in social groups, and have extensive psychological adaptations to social life.

This confusion of levels does conceal several virtues in the target article: The authors are right that a direct mapping of evolutionary theory as a motivational theory is inadequate, and that not everyone recognizes this; their experiments add to our knowledge of the variables that our psychological mechanisms respond to in cooperation; in-group biasing is a phenomenon that must be addressed and explained (we favor a coalitional explanation in which small scale coalitional aggression was likely to have played a prominent role; Alexander 1971; Tooby & Cosmides 1988). Nonetheless, we find the claim that "it was adaptive for ancestral humans to identify automatically with an in-group and to accept its goals as their own" and that "human nature is basically social rather than selfish" as no more plausible a psychological characterization than the idea that humans cognitively represent fitness and selfishly pursue it as a goal. Human nature is not "basically social" or "basically selfish"; human nature is "basically" a collection of mechanisms designed to achieve genetic propagation in our environment of evolutionary adaptedness in ways that are sometimes characterizable as "selfish," sometimes as "altruistic."