

Cognitive Psychology II in Frankfurt: Possible Projects

Sabine Windmann (March 2024)

Altruism from the Perspective of Behavioral Learning Theory

Altruistic acts define social outcome contingencies. They come in three variants: First, **costly rewarding** reflects the sharing of resources to reinforce socially desirable acts and to support individuals who are perceived as needy or deserving. Second, **costly punishment** is exerted to deter defectors from free riding, and to enforce fairness in group cooperation. Third, **costly contingency change** occurs when established reinforcement plans are challenged, against existing norms and power structures, requiring courage and post-conventional morality.

Three theses are derived from this perspective:

1. Altruism is subject to the laws and rules of Behavioral Learning Theory

Findings from behavioral learning theory can be transferred to altruistic acts among humans, and behavioral learning theory can be used to predict the outcomes of social interactions and economic games. Expectancy effects, discount functions, extinction resistance, Kamins Blocking effect, matching law, Premack principle... all these phenomena can be observed in response to altruistic rewarding and punishment. When help is expected but denied, reactance will ensue (negative contrast effect). In a series of social rewards or punishments, the first one will be the most effective. Non-contingent help and punishment will demotivate recipients (learned helplessness). Latent inhibition is behind the Genovese syndrome. People are "biologically prepared" to expect help and punishment, though more so from kin than from non-kin. Conditioned compensatory reaction occurs when costly rewards or costly punishments are given in an unfamiliar social environment, rendering the effects on the ensuing behavior stronger. Predictiveness (absolute and relative) will guide the seeking of help and the avoidance of punishment. Removal of an aversive stimulus can help (= negative reinforcement) whereas removal of privileges and benefits can reduce socially undesired behavior (= negative punishment).

2. Altruism is not a unitary phenomenon: The three altruistic behaviors can be dissociated between and within individuals

Between individuals: There are temporally stable and domain-overarching differences between individuals in their inclinations towards the three altruistic behaviors. Dirty Harry uses punishment, Mother Theresa helps needy others, and Edward Snowden changes outcome contingencies against the threat of powerful others.

Within individuals: Positive mood elevates willingness to reward others, anger increases costly punishment, and a sense of autonomy increases the willingness to change contingencies.

3. The three types of altruistic acts can be reliably assessed in behavior and self-report

The three types of altruistic behaviors are reflected in economic games: Costly rewarding is reflected in the Dictator game, costly punishment in Ultimatum or Public Goods game. A game for contingency change needs yet to be constructed. Behavioral preferences measured with the economic games should correlate with self-reported behaviors.

4. Cross cultural perspective [To be developed]

Ingroup-outgroup dissociations, parochialism, social value orientation as a function of hierarchical structures, institutionalization.