Theories of human altruism: a systematic review

Svetlana Feigin1*, Glynn Owens2 and Felicity Goodyear-Smith2

1Department of Psychology, Faculty of Science, University of Auckland, New Zealand, 2Department of General Practice and Primary Health Care, Faculty of Medicine and Health Science, University of Auckland, New Zealand.

Abstract

The goal of this systematic review was to summarise the social psychological literature on theories of altruism in humans from 1960 to 2014. Material was sourced through online databases, book sections, grey literature and hand searches. All social psychological literature on altruism in humans from 1960 to June 2014 was eligible for inclusion in the review. Items were critically reviewed according to key characteristics and findings. Non-English items and original research papers were excluded. Of 1881 potentially eligible items, 308 were selected for critical appraisal; of these, 97 were included in the review as ‘Theory’ articles, being separated into pseudo-altruism (selfishly-motivated) and true (selfless) altruism. Within these categories lie a range of models. The influences behind altruism are complex and do not arise from a single source but rather a multitude of sources both within and outside the individual. Future theoretical models would greatly benefit from being inclusive rather than exclusive in allowing for the possibility of co-existing motivational drives (egoistic and altruistic).

Keywords: Altruism, altruistic behaviour, empathy, helping behaviour, pro-social behaviour

Introduction

The term ‘altruism’ was introduced by Auguste Comte to describe devotion to the interests of others as an action-guiding principle [1]. For present purposes, the authors defined features of human altruism as an intentional and voluntary act performed to benefit another person as the primary motivation and either without a conscious expectation of reward (altruistic approach) or with the conscious or unconscious expectation of reward (pseudo-altruistic approach) [2-11].

Scientific interest in altruism in humans began to grow from the 1960s onwards [12]. However, conflicting theoretical and empirical literature created several problematic divisions. This literature review presents conflicting theoretical arguments from both the pseudo-altruistic (selfishly motivated) and true (selfless) altruism. Within these categories lie a range of models. The influences behind altruism are complex and do not arise from a single source but rather a multitude of sources both within and outside the individual. The aim of this systematic review was to synthesise the conflicting theoretical arguments and suggest ways to overcome the conflict.

Methods

A systematic literature review was conducted using PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) criteria [13,14] to synthesise social psychological literature on the topic of altruism in humans from 1960 to the present (2014). To narrow the research focus, only social psychology-related papers on the theory of altruism in humans were included, with original research excluded.

Literature was sourced from the major databases (PsycInfo, PsycArticles, PsycCritiques, PsycExtra, Psychology and Behavioural Sciences Collection, ProQuest, Voyager and Google Scholar) as well as secondary sources such as reference lists from accessed articles and grey literature. Sourced literature

How to cite this article:

was initially screened for relevance (title and abstract) and then full texts were assessed using an eligibility checklist.

Results
Of a total of 1881 potentially eligible records, 97 original theory articles were included in the present systematic literature review. Some of the reasons for exclusion of records after full-text review were if the theoretical approaches discussed were non-human specific and were of a philosophical and non-social psychological background. The eligibility criteria used in the full-text review also excluded articles which were largely opinion pieces or commentaries and stemmed from evolutionary arguments. Early evolutionary theories were discussed only briefly to provide the reader with an insight into early theoretical influences on the modern pseudo-altruistic and altruistic division. A summary of the theories of human altruism is presented in Figure 1.

Theories on human altruism
Defining altruism is problematic, with definitions often guiding how research is conducted [15]. One of the most prominent divisions relates to the motivation underlying altruism.

Some argue that to define altruism, one must separate it from its opposite—egoism [16,17]. Accordingly, altruism and egoism are two separate motivational states and are distinct in terms of the direction of goal-directed motivation (to the self or to the other), with altruism's ultimate goal being the increasing of another’s welfare [16,18,19]. While egoism and altruism are potentially distinct motives, they can co-occur [19]. Others argue that altruism is selfishly motivated, with the underlying goal being one’s own well-being [20-22].

Altruism may also be defined as either normative or autonomous [23]. Normative altruism includes common-place acts of helpfulness governed by social rewards and punishments; autonomous altruism is not influenced by these [23]. Examples of autonomous altruism are not commonplace and include acts such as those committed by rescuers of Jews in Nazi-occupied Europe. Such acts may also be referred to as heroism [24]. In general, normative altruism is low risk and low cost for the actor while autonomous altruism is often high risk and high cost.

Researchers presenting altruistic approach arguments agree that the essential features of altruism include an act performed voluntarily and intentionally with the primary goal of benefitting another person [2-11]. The underlying motivation is what separates the altruistic approach from the pseudo-altruistic approach. Namely, whether an act was performed without (altruistic) or with (pseudo-altruistic) the expectation of personal gain whether that be in the form of internal or external rewards. Thus, a critical criterion in the definition of altruism is its motivational base [4,19,25,26].

Early theories
Early evolutionary theories influenced theoretical and empirical...
approaches to explaining human altruism. Specifically, early evolutionary theories overshadowed theoretical understandings of human altruism as possessing underlying selfish motivation. The concept of altruism was something of an anomaly for Charles Darwin’s theory of the evolution of man and natural selection [12] and was largely neglected. It was only in the mid to late 20th century that researchers resolved the paradox of natural selection and altruism by introducing notions of group and kin selection [20], inclusive fitness [12,20,27] and reciprocal altruism [28]. Group and kin selection theories argued that natural selection favoured groups over individuals through an increased likelihood of passing on genes successfully. Inclusive fitness proposed that an individual’s genetic fitness can be measured not only by its survival and the survival of its offspring, but also by the enhancement of the fitness of its kin [20,27]. Reciprocal altruism argued that natural selection favoured altruism even among non-kin because of the long-term benefits [28].

According to these theories, altruism is ultimately selfish because it serves to increase one’s genetic fitness. The development of these biological and evolutionary theories impacted on some early psychological theories concerned with parental care and self-sacrifice [12,29] and influenced psychological theoretical arguments that human altruism is ultimately selfishly motivated.

Many early psychological theories, such as those based on psychoanalysis, were influenced by the belief that all human motivation is inherently selfish or egotistical [16,26,30]. As a result, investigation of altruism was minimal [26]. Thus, Freud (1856-1939) argued that all actions ultimately occur to meet the needs of the self [16]. Human nature is thus hedonistic with the aim of seeking pleasure and avoiding pain [5]. As a result, the existence of altruism was not actively researched for much of the early 20th century [30].

The argument that all human motivation is inherently selfish has long been dominant [31], but the alternative proposal that the existence of unselfish motivation directed towards benefitting the ‘other’ and not the ‘self’ was first referred to under the heading of ‘benevolence’ [32,33]. In 1851, Auguste Comte introduced the term ‘altruism’ to distinguish this form of unselfish motivation from acts which were selfishly motivated [34]. The existence of ‘true’ altruism however, remains contentious and theories on human motivation continue to be characterised by egoistic or altruistic undercurrents.

The pseudo-altruistic approach

The pseudo-altruistic approach has been a dominating force in psychological theory [31]. Its defining feature is that so-called altruistic behaviour is ultimately egoistically-motivated, the end goal being one’s own welfare [26,30,35]. Altruism is thus redefined to fit the argument that all human action is self-serving, differing from Comte’s definition [30].

By this definition altruism is seen to be motivated by the attainment of internal rewards [2,7,8,16,30] even when these rewards are not directly observable [36]. Some argue that through an internalisation process, by adulthood altruism acts as a self-reward mechanism [37]. This mechanism eventually fosters self-satisfaction and self-esteem in adolescence [8,30,36,38-40]. This is not to diminish the importance of vicarious empathic emotion, seen as a major criterion of altruistic behaviour [36,41,42]. If however the person develops a reliance on material rewards then the internalisation of intrinsic motivations to help may not be very successful [43]. This view is probably the most popular among those psychologists who adhere to the pseudo-altruistic argument [16,30].

Social learning perspective

The social learning perspective argues that our moral responses are acquired through the ‘laws of learning’ [44]. Internalisation of values is facilitated by observational learning [15,36,40,45]. Parental models have been proposed to exert the strongest and most prolonged influences on the internalisation process [45]. Some theorists also propose that behaviour-contingent learning influences the development of altruistic behaviour during childhood and this in turn serves as an internal reinforcement [36]. Others argue however that social learning does not explain altruism entirely, ignoring genetic influences and high-order reasoning such as role-taking abilities, and thus that social influence is insufficient in explaining moral reasoning [46].

Normative theory

According to normative theory, there are three basic influences on altruism - the intensity of moral (personal) obligation, a cognitive structure of norms and values, and the relevance or appropriateness of feelings of moral obligation [47]. Moral personal obligations or norms are influenced by shared group expectations about appropriate behaviour and social rewards, varying from individual to individual [39,47,48]. People help because they perceive it as the appropriate social response either due to previous experience or observation of others [45]. In essence, people are socialised in society to adopt the norm of social responsibility and help others [9,49]. Related to this are fairness and the need to see the world as ‘just’. These are central concepts to the ‘just world hypothesis’ proposed by Lerner [50-52]. According to this hypothesis there is a shared belief that the world is fair and people get what they deserve and vice versa. Thus, we help those who have helped us and not those who have denied us help [53].

Personal norms also play an important influential role both cognitively and affectively [45]. For example, people may possess expectations of behaviour based on personal standards or experience emotions (such as guilt) when meeting or not meeting these [45].

Stage theoretic approach

The stage theoretic approach argues that individuals go through several stages of ethical development [54] with the central source of change being cognitive-structural development [55]. Early ethical development is mainly rooted in self-interest [56]. Some then develop the motivation to follow community
norms. Others, whose ethical development progresses, acquire universal moral norms [56]. Human behaviour is not however solely dependent upon stages but involves an interaction with situational influences [55]. A truly altruistic act is a developmental achievement, possible only in the final developmental stage or with increasing maturity [8,55].

**Arousal-reduction and negative state relief models**

The arousal-reduction and negative state relief models argue that altruism is motivated by reduction in aversive arousal or tension [16,22,30,57-61]. When an observer witnesses another’s suffering this causes negative emotional arousal which is reduced by helping the sufferer [45,57,58,62]. The experience of guilt in response to another’s suffering can be a powerful motivator [31,63-65]. Feelings of guilt endanger self-esteem and helping behaviour enables self-image reparation [63]. Other events that make the observer feel better can diminish the motivation to help, especially if these events precede the opportunity to help or if the observer is aware of less-costly ways to improve their mood [45,61,66]. Thus, motivation is dependent upon helping possessing self-rewarding properties [22,60,67]. If, however, helping is unlikely to relieve negative emotions, the observer is unlikely to help [68].

**Cost-reward model**

The cost-reward model encompasses costs and rewards for helping or not helping and relates closely to arousal-reduction models [57,58]. The cost-reward model proposes that witnessing the distress of another creates unpleasant empathic arousal in the observer and the observer is thus, motivated to reduce it [57,58]. Empathic arousal motivates the bystander to act and the cost-reward analysis provides direction for the observer’s actions [45].

Slight variations on this model exist. Hoffman [20,31] proposed that witnessing another’s suffering causes empathic distress in the observer and that attributions regarding causes of the suffering influence the observer’s willingness to help [69,70]. Suffering by members of our ‘in-group’ causes significant aversive tension and the perception of a discrepancy between the current and desirable state of another’s welfare produces cognitive inconsistency [71,72].

**Decision model of bystander intervention**

The decision model of bystander intervention [73,74] proposes that the decision to help depends on five factors. The first and second are the identification of a negative change in the victim’s circumstance and recognition that help is needed [73]. The third and fourth factors are the taking of personal responsibility and the decision as to what kind of help is needed [73]. The final factor in the decision-making process is the decision to provide help [73]. If, at any stage the bystander fails to attend to any of the factors, the victim will not be helped [73]. External influences such as aspects of the physical and social environment can influence whether a situation is interpreted as negative [75].

For example, researchers argue that helpfulness is lower in cities than in towns as a result of environmental input overload [76]. Internal emotional states of the bystander can also influence their decision to help. Thus, it has been argued that positive internal mood states are more facilitating for helping behaviour than negative mood states [63].

In addition to these influences, bystanders are affected by those around them. For example, when the bystander is one among several others, diffusion of responsibility may occur [45,77,78]. This refers to the belief that others present are just as capable of helping the victim [77]. Diffusion of responsibility is more likely to occur when an element of danger is involved, others are perceived as being able to help, and/or when norms permit [57,79]. Audience inhibition can also occur, whereby the bystander may be inhibited by fear of embarrassment or negative evaluation [78].

**The altruistic approach**

The second category of contemporary theory on human motivation—the altruistic approach—retains the meaning of ‘true’ altruism as Comte intended [30]. Here, motivation is directed toward the end goal of increasing another’s welfare and any feelings of self-reward or alleviation of personal distress are by-products of this [26]. By challenging presumptions of individualism and egoism, ‘true’ altruism is no longer impossible but rather, ubiquitous [35]. The existence of a ‘trait’ of altruism and of an ‘altruistic personality’ type becomes possible [2,42,80].

Underlying the ‘altruistic personality’ are motivations such as empathy, norms of appropriate behaviour, and a tendency to experience cognitive and affective empathy [42,80]. Role-taking abilities provide the capacity to empathise and norms of behaviour are internalised, guiding judgements [80]. These motivations are however, ‘hypothetical constructs’ [80]. Nonetheless, proponents of the ‘altruistic personality’ describe an altruist as someone with higher standards of justice, social responsibility, modes of moral reasoning, who is more empathic to the feelings of others [80].

**Empathy**

Empathy involves feelings of sympathy and a desire to relieve another’s suffering [81]. Its cognitive component is dependent upon the developmental stage of the observer. As children pass through developmental stages, they eventually become aware of others having personal identities, and are more likely to experience empathy as subjective as well as involuntary arousal [31,82]. In essence, an intrinsic altruistic motive system exists in which a person’s empathic response to another’s distress, coupled with a cognitive sense of the other, provides a basis for a motive independent of egoistic motivation [83].

From these components of empathic distress emerges sympathetic distress which is argued to be the underlying motive for altruism [31,81,82]. Empathic distress may have the opposite effect, however, directing the observer’s attention away from the victim onto themselves, and decreasing the likelihood of altruistic behaviour [31,82]. This suggests a possible optimal range of
empathic arousal for altruistic behaviour. Observers may also try to reduce the level of arousal, employ different strategies (such as looking away) or derogating the victim [31,82,84]. Furthermore, if victims are perceived to be responsible for their own plight the empathic distress may be neutralised [31]. Thus, empathic and sympathetic distress is pro-social only within certain limits.

**Autonomous altruism**

Specific emotions have been proposed as triggers of altruistic motivation [23,81,85]. According to this view, whilst experiencing these emotions, the observer violates laws of reinforcement by directing behaviour towards the needs of others without regard for oneself [85,86]. Autonomous altruism is an example of this [23]. By definition, autonomous altruism should be seen as ‘selfless’ as it is not governed by societal norms and its impetus comes from the self. Furthermore, a process of ‘empathic conditioning’ has been said to occur in both children and adults whereby the observer’s behaviour is influenced by particular cues which elicit an empathic response [81,87].

Karylowski [59] argued for a distinction between doing good to feel good about oneself (endocentric altruism) and doing good to make another person feel good (exocentric altruism). According to this view, endocentric altruism is pseudo-altruistic while exocentric altruism is purely altruistic in motivation [30,59]. Exocentric altruism is argued to be a result of the focusing of attention on the other rather than the self [59].

**Just world hypothesis**

The just world hypothesis was expanded by Lerner to include identity relation [88]. When we perceive ourselves as ‘psychologically indistinguishable’ from another we experience what they experience [88]. When we perceive the other person as in need, ‘justice of need’ is evoked, and we engage in ‘identity-based activities’ [88]. The self becomes indistinguishable from the ‘other’ and the underlying motivation to help cannot be exclusively egoistic or altruistic [30,88]. This view is similar to that of Hornstein’s [72]. In particular, the similarity is evoked through the experience of ‘we-ness’ [30,72] and the dissolution of the self–other distinction [30,88].

**Empathy-altruism hypothesis**

A common theme in altruism theories is that empathic emotion produces altruistic motivation to help [16,30,36,65,83,89–93]. Batson, describing three key paths leading to helping includes one associated with empathically evoked altruistic motivation [16,30], referred to as the empathy-altruism hypothesis [16,18,26,92–95]. This path is of particular interest to the current literature review separating itself from the dominant egoistic argument of underlying motivation for helping. It is characterised by seven empirically testable processes [16,30]. The first, ‘perception of another in need’, is a function of several factors [16,30] including a perceptible discrepancy between the other’s current and potential state of well-being; sufficient salience of these states, and focus of attention on the other [16,30]. The thresholds for magnitude include the number of dimensions of well-being perceived to be discrepant; the perceived size of the discrepancies, and their perceived importance [16,30]. The perception of the other’s need leads to empathic emotion [16,30].

The second process involves the adoption of the other’s perspective and is a function of the threshold of two factors: ability to adopt the other’s perspective [25,31,65] and a perspective-taking set [96]. Attachment is the third psychological process and contributes to the experience of empathy, which is the fourth psychological process [16,30]. Arousal of empathy is affected by attachment in two main ways. Specifically, the strength of attachment affects the likelihood of adopting the other’s perspective and the strength of attachment can affect the magnitude of the empathic emotion experienced by the observer [16,30]. Unlike the experience of personal distress which may evoke egoistic motivation, the experience of empathy, characterised by feelings of sympathy and compassion, evokes altruistic motivation (fifth process), the ultimate goal of which is the improvement of the other’s welfare [16,30]. The magnitude of altruistic motivation (fifth process) is dependent upon the magnitude of the experience of empathy [16,30,92]. Altruistic motivation may elicit social and self-reward as well as avoidance of punishment and reduction of personal distress which form the basis of egoistic motivation. Batson [16,30] argues however, that these consequences are only a by-product of the motivation and not its ultimate goals. Furthermore, according to the empathy-altruism hypothesis, ‘true’ altruism does exist and is a product of psychological processes.

The sixth psychological process, termed ‘hedonic to calculus’ or relative benefit analysis, is an analysis performed by the observer with the intention of determining the most effective behaviour possible to attend to the needs of the other and whether someone else is more capable of achieving this [16,30]. It can be argued that this is egoistic in nature; however this should not be interpreted as meaning that the motivation to help is also egoistic. Furthermore, the altruistically motivated observer will help so long as helping is possible and the relative benefit analysis is in favour of this [16,30]. In cases when the relative benefit analysis is negative, the observer will avoid helping by either simply ignoring the victim or derogating the victim [51]. In this case, the empathic emotion along with the altruistic motivation is wiped out.

The empathy-altruism hypothesis has however, been criticised for attempting to argue the existence of ‘true’ altruistic motivation and instead, it has been suggested that the underlying motivation is egoistic. For example, some researchers have argued that manipulations used to elicit empathy in participants also elicit a self–other overlap or a sense of ‘oneness’ which contributes to an improvement in the psychological well-being of the observer once help has been given to the person in need [97]. Batson [98] however, presented findings suggesting that group membership or the experience of ‘oneness’ does not affect helping behaviour. Batson and Coke [26] also note that providing opportunities of easy escape to the observer should
Altruistic theories suffer from several problems. Firstly, they lack precision and are sometimes misinterpreted [30]. Secondly, a lack of precision in their definition means that empirical findings are often inconclusive and can be accounted for by pseudo-altruistic theories [30]. Nonetheless, altruistic theories provide new interpretations and theoretical arguments to the study of altruistic behaviour which separate themselves from selfishly-motivated explanations.

Furthermore, it is particularly difficult to make concrete conclusions about the existence of what has been referred to as ‘true’ altruism in humans. The dense theoretical literature on human altruism stems from many different theoretical viewpoints. Clearly the motivational (intra and interpersonal) and behavioural influences behind altruism are complex in nature and do not arise from a single source but rather a multitude of sources both within and outside the individual. Existing theories highlight many elements and provide some insight into human altruism. From existing literature, we can begin to develop empirical investigations and theories and perhaps realise the complexity of the construct of altruism.

Certainly it is clear that human egoism is insufficient in explaining all aspects of ‘altruistic’ behaviour. It thus follows that the assumption of universal egoism must be replaced by a more complex assumption allowing room for both egoism and altruism [19].

Some intrinsic limitations of the present literature review existed. For example, non-English language articles were excluded which may have limited the results to only English-language articles. Additionally, as the database search was done online, older articles not included in online databases may have been missed. Finally, the present literature review is only comprehensive of social psychological theories on human altruism and thus, does not cover theories from other psychological disciplines such as organisational and evolutionary.

Future directions
The complex and multi-faceted nature of humans seems to suggest a need to be ‘open-minded’ when developing theoretical models. Lack of an ‘open mind’ in relation to understanding the underlying motivational forces of human altruism will only serve to handicap future research and theories. Future theoretical models would greatly benefit from being inclusive rather than exclusive in allowing for the possibility of co-existing motivational drives (egoistic and altruistic). The conflict between pseudo-altruistic and altruistic theoretical approaches can be minimised or resolved by combining elements from both sides of the debate and conducting research exploring the validity of these new combinations.

Competing interests
The authors declare that they have no competing interests.

Authors’ contributions
All authors contributed equally to the manuscript.

References


50. Lerner, MJ. (1980). The belief in a just world: a fundamental


Feigin et al. 2014