Autonomy and morality: A Self-Determination Theory discussion of ethics

Alexios Arvanitis
University of Crete

NOTICE: This is the author’s version of a work that was accepted for publication in NEW IDEAS IN PSYCHOLOGY. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in NEW IDEAS IN PSYCHOLOGY, [Vol. 47, pp. 57-61, December 2017] doi: 10.1016/j.newideapsych.2017.06.001
Now online at https://authors.elsevier.com/a/1VBa5.GfNqNHHv (free access until 30th July 2017)

Author Note
Alexios Arvanitis, Department of Psychology, University of Crete.

I thank Marietta Papadatou-Pastou, Konstantinos Kalliris, Edward Deci and Richard Ryan for helpful comments on an earlier version of this paper.

Correspondence concerning this article should be addressed to Alexios Arvanitis, Department of Psychology, University of Crete, Rethymnon, Greece. Tel.:+306973329852. E-mail: a.arvanitis@uoc.gr.
Abstract

Kantian ethics is based on a metaphysical conception of autonomy that may seem difficult to reconcile with the empirically-based science of psychology. I argue that, although not formally developed, a Self-Determination Theory (SDT) perspective of ethics can broaden the field of Kantian-based moral psychology and specify what it means, motivationally, to have autonomy in the application of a moral norm. More specifically, I argue that this is possible when a moral norm is fully endorsed by the self through a process of internalization that is energized by intrinsic motivation and is facilitated by the fulfillment of the basic needs for autonomy, competence and relatedness. The conditions of the fulfillment of these needs may even implicitly reveal which norms will be integrated and treated as moral norms. I conclude that SDT offers a motivational approach that is useful in understanding the development of moral norms.

Keywords: ethics; morality; motivation; autonomy; self-determination theory; Kant
Autonomy and morality: A Self-Determination Theory discussion of ethics

Moral psychology is an interdisciplinary field that draws from psychology and philosophy equally and cannot really be approached without a basic understanding of both. This becomes apparent when one gets acquainted with the best known theory of moral psychology, that of Kohlberg (1971, 1973). Kohlberg defined the highest form of moral development in terms of the theory of Immanuel Kant (1785/2011) and conceptualized it as our ability to author the moral rules we live by. Kant (1785/2011) saw ‘autonomy of the will’ as the foundation for authoring our own rules and famously argued that ethical principles should not be derived from human inclination: What we ought to do should not stem from what we want to do. Our needs, our emotions, our intuitions should be sacrificed in favor of objective criteria such as the categorical imperative, that is, acting consistently with a rule that we would treat as a universal law. This ability entails a high level of freedom from environmental contingencies as well as self-serving predispositions. Under this perspective, neither rewards nor conventional rules can be the basis of morality. It is through our own reasoning that we transcend our own bodies, author rules as universal laws and achieve the highest level of morality.

Kant’s theory leaves little room for psychology, since his view is based on a metaphysical conception of will and duty that have no empirical antecedents (Campbell & Christopher, 1996). Any type of empirical cause would undermine the autonomy of the will and would taint the moral character of an act. It is not easy to take his theory and apply it as a guide for approaching moral behavior in the real, empirical world. This may be one of the most important reasons for why Kohlberg’s approach seems rationalistic: it
depends too much on the abstract interpretative scheme of Kantian reasoning and on a hierarchically ordered sequence of moral cognitive structures that are more logically than psychologically related to each other. This fact may also account for why his perspective is no longer dominant. More recent approaches incorporate factors besides rational reasoning in their analyses, such as emotions (Eisenberg, 2000; Hoffman, 2001), skills (Churchland, 1998), virtues (Peterson & Seligman, 2004) and intuitions (Haidt, 2001). Scientists are exploring new ground and including affective or personality processes, but also societal and cultural influences, in order to explain moral judgment. Of course, some approaches rely mainly on social processes (e.g., Ellemers & Bos, 2012) whereas others may rely on biological factors, such as brain structure (Moll, Zahn, de Oliveira-Souza, Krueger, & Grafman, 2005).

Even though Kant’s theory of freedom has received considerable criticism --even within philosophy--, it does make a strong, simple point that is easily, intuitively understood: If my reasons for performing a seemingly moral act illustrate contingencies beyond the act itself, that act should not be considered moral. For moral psychology the next question to ask is how we can discuss this type of autonomous behavior without stripping individuals of their brains, their emotions, their skills, their goals. In other words, without ignoring their organisms.

A potential answer may lie with a motivational theory called Self-Determination Theory (SDT), which examines the reasons why people perform any type of activity and places great importance on autonomy. Autonomy is defined as regulation by the self (Ryan & Deci, 2006) and is distinguished from heteronomy, that is, regulation by forces outside the self. One of the most
important aspects that SDT can bring into the discussion of moral psychology is that it also distinguishes autonomy from independence (Ryan, 1993), meaning that it defines autonomous acts not in terms of the pure absence of contingencies but in terms of a person’s endorsement of the act itself. The question I will ponder on is whether true endorsement of a moral act, even in the presence of contingencies, succeeds in the maintenance of the character of the Kantian moral act. In this endeavor, I will focus on autonomy and the concept of internalization, discuss SDT’s relation to other theories, and argue that the integration of ethical rules is the basis of true, autonomous morality.

A motivational account of ethics

A motivational approach would focus on why an individual is energized or activated toward a particular end or, in the case of morality, why a person is energized toward ethically appropriate behavior. Motivation is closely related to the notion of energy, discussed in the classic theories of Freud (1962) and Hull (1943) and relating to the notions of instincts and drives. Especially after the advent of behaviorism, these theories declined in significance because of their inability to account for normal development (within Freudian theory) and to account for exploratory and interest-based behaviors (within Hullian theory). In response to the decline of motivation as a topic for research, White (1959) shifted the focus from the concepts of instincts and drives to independent ego energy and intrinsic motivation. He argued that exploration, playful behavior, and the production of effective changes in the environment exhibit direction, selectivity and persistence in a way that the actions seem motivating in their own right. Moreover, these aspects of behavior can only be conceived in relation to the stimulus field that an
individual would aspire to “conquer” and, therefore, concentration on cognition and emotion alone, as is often the case, cannot fully account for human behavior. Similarly, it can be argued that morality should not be sought in isolated aspects of human behavior—such as rational reasoning or emotions— but should take into consideration a broader conception of behavior that takes into account these inherent motivational tendencies.

Self-Determination Theory relies on a quite similar premise to White’s (1959) propositions: individuals exhibit intrinsic motivation as the manifestation of the built-in energy of the organism to seek and conquer optimal challenges (Deci & Ryan, 1985). An SDT view of ethics will therefore attempt a more holistic psychological account of ethical behavior since it does not emphasize isolated elements of the person or the environment but the interplay between the two. On the one hand there is the environment, aspects of which often seek to control behavior, and on the other hand there is the organism, seeking to expand and conquer aspects of the environment. Self-Determination Theory accepts that the environment can control behavior but also that individual decisions can be causal elements in behavior. In this interaction, it is the will that energizes support for individual decisions and allows a person to counteract forces from the environment (Deci, 1980). In fact, the notions of will and autonomy are central elements in Self-Determination Theory, just as they are in Kant’s account of morality. With regard to morality, an SDT account would especially focus on exploring when the will energizes support for ethical decisions.

The act, the self, contingencies and autonomy
In the Groundwork of the Metaphysics of Morals, Kant (1785/2011) argued that the good will is not defined as good by what it accomplishes; it is good in itself. It is not defined, for example, by the fact that it may guide actions that contribute to the happiness of others. It is not defined by any sort of inclinations or natural laws either. In fact, the “good” will is not defined by anything other than reason. The “good” will is purely autonomous, free from contingencies and inclinations, and will abide by rules out of duty alone. Duty is a central element of Kantian ethics and refers to the objective necessity of an action from obligation. Of course, scientists struggle with such a metaphysical account of duty and will.

For an empirical science, such as psychology, it is indeed difficult to accept that the so-called *noumenal* will, having no connection to the empirical world, can actually cause behavior (Campbell & Christopher, 1996). Kant’s concept of transcendental freedom goes beyond the organism and any contingencies of the empirical world into the realm of reason where the person can truly be free. This is the field where moral norms are produced according to Kant, who evidently wants to disengage the production of moral laws from any sort of environmental influence. Any type of duty should arise autonomously, free from contingencies. From a psychological perspective, I will try to account for how an organism can act out of duty while minimizing any sort of contingencies. In this sense, I offer a Kantian-based psychological account of ethics without accepting its metaphysical underpinnings.

Self-Determination Theory argues that an autonomous act is defined as regulation by the self, the *self* being a central process that regulates behavior and experience. It is an organismic theory that accepts that humans have a
natural propensity to grow and assimilate aspects of their environment. Behavior is essentially the product of the interaction between the organism and the environment. A self-determined, autonomous act is an act that is regulated by the phenomenal ego-center (Pfander, 1908/1967), whereas a heteronomous act is mainly caused by external influences such as environmental controls or even inner impulses that are experienced as controlling urges that are not endorsed by the phenomenal ego-center or self. It is not the absence of controls that makes an action autonomous but full endorsement of an act by the self, even if it is in accord with an external demand (Ryan & Deci, 2006). Within SDT, there is a concept that reflects the perceived autonomy that the self experiences in the performance of an act: The Perceived Locus of Causality (PLOC- Ryan & Connell, 1989; cf. Heider, 1958). Acts with an internal PLOC are acts in which a person feels as the origin of the behavior whereas acts with an external PLOC are acts in which the person feels as a pawn (cf. De Charms, 1968). Autonomous acts are therefore those acts during which the person feels as the “owner” of the behavior, even if they are in line with external demands. Extending this reasoning to the field of moral psychology, autonomous moral acts would be moral acts in which the individual perceives oneself as the origin of behavior, irrespective of the existence of environmental controls and inner impulses.

The prototype of an autonomous act is an intrinsically motivated activity where a person performs the activity for the sake of the activity itself (Ryan & Deci, 2000). In this case, there is no “ought” prevalent in the situation, no external constraints, but only wants and inclinations--such as curiosity, exploratory or playful tendencies--that are satisfied during the act.
Intrinsic motivation is the primary energizing force within the organism that strives for the engagement in one’s interests and the exercise of one’s capacities (Deci & Ryan, 1985). It exists within the individual--and not necessarily only in humans--although it is understandable that the focus can also be on inherently interesting properties of an activity (Ryan & Deci, 2000). By Kantian standards, intrinsically motivated behavior cannot be considered moral because the main connection between the activity and the self is an inclination and not a sense of duty. No act that appears to satisfy inclinations can be considered the basis of a moral norm.

Indeed, what we are in search of is an extrinsic connection between the activity and the self that takes the form of a moral norm, imposes a certain sense of “ought” but does not appear to embody any contingencies that transcend the activity itself. Although the prerequisite of the existence of an “ought” standard excludes the possibility that a moral act can be an intrinsically motivated act, the autonomy component of intrinsically motivated behavior is a sought after property for true moral behavior because there is no instrumentality, there are no external contingencies guiding behavior. This is the type of Kantian autonomous property that we are searching for in a moral act: Although the act itself is not the purpose, there are no contingencies present in the individual’s motivation. The primary motivation of a moral act should therefore exist within the person’s will to proceed with the act and should not be grounded in any type of contingencies, even if they are objectively present in the situation.

Let us take the example of a person diving into the water in order to help a man who is drowning reach the shore. If the person is a competitive
swimmer and jumps into the water only because he loves how the cold water feels and enjoys performing under challenging circumstances, then the activity can be described as intrinsically motivating because the act itself is the purpose. It is motivated by inclinations, exploratory and playful tendencies. If the purpose, however, is to save the man’s life, then this act is not intrinsically motivated. On the contrary, it is extrinsically motivated because it serves a different purpose than the performance of the act itself. In other words, it is a means to an end. Moreover, if the act does not appear to serve a higher ideal, a moral norm, it is doubtful that the act will be described as moral. However, not all extrinsically motivated acts that appear to comply with a moral norm will be treated as moral. According to Kant, they should be autonomous. Therefore, under a Kantian-based view of ethics, the nature of the reasons for the moral act should also be examined.

According to SDT, there are four types of extrinsic motivation (Ryan & Deci, 2000): External regulation, where the act is performed for the attainment of certain rewards or the avoidance of certain punishments; introjected regulation, where the act is performed in order to protect or boost one’s own self-esteem, avoid guilt or feel pride; identified regulation, where a person realizes the significance and value of the act; integrated regulation, where the act is fully in tune and consistent with the totality of a person’s identifications and value system. To refer back to the previous example, if there was a high reward to save a man’s life and this was the reason why the person dove into the water, the rescue would be classified as an externally regulated act. If the person dove into the water to prove to himself how brave he is, it would be classified under introjected regulation. If the person dove into the water on the
belief that every human life is valuable, this act would be classified under identified regulation. Finally, if the person dove into the water despite the realization that his life would be jeopardized and his loved ones distressed, taking into account all relevant factors but still endorsing the principle of what he ought to do in this kind of situation, the act of diving into the water would be classified as integrated regulation.

The more the extrinsic properties of the situation appear close to the fundamental phenomenal core of the person we call the self, the more the reasons for performing an activity are difficult to dissociate from the self or the activity itself. Contingencies may be objectively present, but they do not define the person’s motives for performing the act. With regard to the previous example of the dangerous endeavor of trying to save a person’s life, even the morality of an identified regulated act might be brought into question if the prospective rescuer entirely neglected to think about his own children. This thought could reverse his decision by making the possible effect on the children’s happiness salient, in case something went dreadfully wrong. On the other hand, an integrated regulated act already incorporates the consideration of similar contingencies, by bringing all identifications in harmony with the person’s unified sense of self—through the process of internalization and its last stage, integration, which will be analyzed further--. This is why integrated regulation seems to be very close to intrinsic motivation (Ryan & Deci, 2000): It is so natural an extension of the self that all contingencies are absorbed into the person’s unified sense of self. This is also why external regulation and introjection are considered heteronomous types of motivation whereas regulation through identification and, more so, integrated regulation are
considered more autonomous types of motivation. Finally, this is why under Kantian ethical standards we would treat an integrated regulated act which is consistent with a moral norm, as a true moral act: It is in line with a moral norm while at the same time, any present contingencies do not define the person’s act. The act is a natural extension of the unified sense of self.

**The internalization of moral norms**

It is difficult to imagine how acts can become natural extensions of the self. The move from external regulation to integrated regulation is a complicated process that is called *internalization*. According to SDT, internalization is propelled by intrinsic motivation, the active process of pursuing one’s own propensities and developing one’s own capacities (Deci & Ryan, 1985). It is not a passive one-way process directed from the environment to the organism but an active accommodation of the environment by the organism. Therefore, the process of moral development, that is, the move toward integrated moral norms, is driven by the fundamental aspects of intrinsic motivation. I should note here that researchers in SDT have not explicitly studied the development of moral norms; they have studied behavior in general. I will, however, use relevant research and make predictions concerning moral norms. At the same time, I have to acknowledge that these predictions should be tested by empirical research, as SDT is an empirically-based theory.

**Self-Determination Theory** argues that there are three basic psychological needs for competence, relatedness and autonomy that display the deep structure of human psyche and reflect the basic tendencies for effectiveness, connectedness and coherence (Deci & Ryan, 2000). These basic needs reflect the energizing force of the organism and can provide a roadmap
of how moral norms can be integrated. As long as these needs are not thwarted, SDT argues that internalization of regulations and values will be a natural process for the healthy functioning organism.

Regulations and values are more easily internalized when the environment provides the nutrients for the support of the needs for autonomy, competence and relatedness (Deci & Ryan, 2000). From the autonomy side, an environment where contingencies are not prevalent will help individuals actively seek and internalize the moral norms that express their natural tendencies. This is possible when rewards are not salient and guilt or anxiety fostered by circumstances such as ego-involvement (Ryan, 1982) is avoided.

From the competence side, the existence of ethical challenges that a person can successfully live up to--in contrast to ethical challenges that a person will find extremely difficult to fulfill--will bring out moral norms that will be more easily internalized. Finally, from the relatedness side, an environment where a person has the ability to connect to others and socialize will foster the internalization of relevant moral norms.

As a general rule, an environment promotes a self-determined adoption of a moral norm that appears in the social context through the existence of low contingencies, optimal challenges and the chance to connect to others. Deci, Eghrari, Patrick, & Leone (1994) elaborate especially on the processes of integration—which could be considered as the last step of internalization--and provide a framework that seems pertinent to the integration of moral norms. It would not be a big leap to generalize their research into the realm of moral development and argue that three contextual events will allow integration: (a) providing a meaningful rationale that will help a person understand the
significance of a moral norm (b) acknowledging the feelings of conflict between an inclination and the fulfillment of the moral norm, (c) conveying choice so that the person will be allowed to take responsibility.

The concept of moral internalization is also present in the classic work of Freud and Piaget. Freud uses concepts such as introjection (Freud, 1924) or identification (Bronfenbrenner, 1960)—defined differently within SDT—but focuses on the role of the parent and sexual instincts as the primary energizing force of internalization. Morality is restricted to the superego and is in constant opposition with the id, reflecting the aforementioned conflict between inclination and morality. The solution to this conflict resides in the synthetic function of the ego (Freud, 1962). On the other hand, for Piaget, two types of morality exist: A heteronomous type of morality of constraint that imposes a sense of duty, characterized by unilateral respect toward authority figures, and an autonomous type of morality, based on reciprocity and justice (Piaget, 1948). According to Piaget, socialization and cognitive development will favor the autonomous type of morality.

According to SDT, though, the move toward more autonomous, contingent-free morality constitutes a move from externally regulated or introjected moral acts, which are associated with rewards and punishment, to moral acts that are supported by integrated regulation. This move is made possible by the intrinsically motivated process of an organism toward self-determination. In contrast to Freud, SDT does not focus on instincts as the primary energizing force of internalization whereas in contrast to Piaget, it offers no assurance that a person will in fact move toward integrated forms of
morality. The process of internalization is based on the support of the basic psychological needs for autonomy, competence and relatedness.

Self-Determination Theory should also be compared with Kohlberg’s theory of moral development. Kohlberg opposes internalization theories because they seem to focus on relativistic cultural conceptions of morality rather than universal principles of moral development (Kohlberg, 1971). However, this opposition is not very strict and there may be ground for the reconciliation of his theory with theories that focus on society and affect (Gibbs, 2003; Gibbs & Schnell, 1985). Kohlberg focuses on a structured sequential set of six stages that the individual reaches through the development of logic and reasoning. The stages are divided in three main levels: Preconventional, conventional, postconventional. The preconventional level focuses mostly on rewards and punishments, in a similar vein to SDT’s external regulation. The conventional level is derived from the power of conventions to impose authority and conformity which SDT would mostly classify under introjected regulation. The postconventional stage, which Kohlberg treats as autonomous, reflects an effort to abstract from specific contexts and create broader, even universal, principles. This postconventional stage is similar to SDT’s integration level.

Despite some similarities, there are major differences between the two theories, the most important of which concern how people move from one stage to the next. Kohlberg argues that every level reflects a general moral state of the person that concerns all moral behaviors and regulations and can only be achieved step-by-step; SDT treats every regulation and behavior separately and situations are not treated uniformly according to some type of
general moral level. According to SDT, it is possible for one behavior to be externally regulated and a similar one to be regulated through identification. Moreover, Kohlberg’s theory is based on an individual sequential development of reasoning whereas SDT is based on the internalization process that reflects the interaction between individual and environment. It is not therefore driven by simple cognitive development but by intrinsic motivation and the environment’s support or thwarting of psychological needs.

**Integration and moral relativity**

According to SDT, a person is free—or not—in the empirical world during the person’s interaction with the environment. The point of convergence between SDT and Kantian-based ethics concerns the freedom from contingencies during the adoption of a moral norm. I have argued so far that freedom from contingencies is motivationally possible only on the level of integrated regulation.

The main question that arises concerns the universal or relational character of the norms that are integrated. Does the fact that people have their own personal and often different inherent tendencies mean that the intrinsically motivated process of integration would produce moral norms that are tailored to the individual? If so, SDT is a long way from Kantian morality because the latter focuses on universal moral norms. However, the fact that psychological needs seem to exist cross-culturally (Chirkov, 2009; Chirkov, Ryan, & Willness, 2005) suggests that the *procedure* of internalization is the same universally. If the procedure is the same universally, maybe the content of morality will exhibit some basic, universal attributes.
A philosophical theory that has expanded Kantian ethics and also focuses on a certain procedure--that of moral argumentation--is Habermas’s theory of discourse ethics (Habermas, 1991). Habermas’s discourse ethics continues a tradition of theories that focus on the procedure through which moral norms are formed. Other famous theories include Rawls’s “original position” where parties determine the fairness of a situation behind a “veil of ignorance” (Rawls, 1971) and Mead’s concept of ideal role-taking (Mead, 1934). Habermas (1991) argues that unless people participate in interpersonal relations, they are unable to form their identity. This is why, intuitively, both the individual and society should be protected at the same time. The ideas of justice--as the equal respect of individual rights--and solidarity--that postulates empathy and concern for one’s neighbor—reflect the interdependence between individuals and their community. Habermas argues that these basic ideas are inherent in the relations of symmetry and reciprocity that are presupposed in his theory of communicative action (Habermas, 1985) and arise during communication. The emphasis of his moral philosophical approach is on the procedure of development of moral norms rather than identification of their content and Habermas concedes that the latter is the object of study for the historical and social sciences.

An SDT-based psychological process of internalization may also provide a procedural account of the development of moral norms while, at the same time, offering the ground for the study of the content of moral norms. The mechanics of organismic integration and its underpinnings can offer us some basic principles for the type of moral norms that will be integrated: It will simply be those norms that facilitate the psychological process of
integration. If a society is structured around ideas that provide the necessary nutrients for the psychological process of integration, it is likely that the ideas themselves will be integrated. In fact, the ideas of justice and solidarity are likely, indeed, to facilitate integration since justice protects the needs of competence and autonomy whereas solidarity protects the need for relatedness.

If environmental moral norms respect the individual and protect the web of interpersonal relationships, the needs for autonomy, competence and relatedness are likely to be supported also. It is then expected that aspects of the environment, including the very principles around which the environment is built, will be internalized, or even integrated. In a similar manner to Habermas (1991), SDT may predict that the principles of justice and solidarity are closely related to the process of moral development. In this case, however, the emphasis is not socio-structural but psychological.

In order, though, to fully examine which specific moral norms will prevail, we need to understand which moral norms, if incorporated into the very structure of the environment, are likely to result in the fulfillment of psychological needs and are therefore expected to be integrated more easily. Although I have argued that the basic form of these moral norms will be similar to the nature of moral norms proposed by philosophical theories, I should stress that psychological research treats these philosophical theories as useful material for the production of hypotheses. It will be the objective of future empirical research to examine the exact nature of the moral norms that will facilitate the fulfillment of psychological needs and will foster integration. This can be a new and exciting field for empirical psychological research.
On a last note, an important question to ask concerns the heterogeneity of moral norms found across different societies, moral norms as diverse as those of “Nazis and Quakers, head-hunters and Jains” (Haidt & Joseph, 2004, p.55). According to SDT, some social norms do not facilitate integration and will only be maintained under the presence of rewards, the threat of punishments or the introjected power to protect and enhance self-esteem. These social norms cannot be considered as true moral norms because they are heteronomously maintained on the basis of particular contingencies. Social norms of this kind may even point to violence or crime. Ryan and Deci (2017, p.641) argue that “malevolent crimes and social practices that harm innocent parties are, in fact, not typically autonomously motivated or motivated through integrated regulations, as would be required definitionally for acting with virtue or morality”. On the other hand, some social norms will facilitate integration and will be endorsed by individuals autonomously: These are the norms that have the potential to create the context for psychological need fulfillment and will be celebrated as true moral norms. A formal development of a motivational theory of ethics will allow us to evaluate and understand not only the effect of moral norms on the individual, but also the morality of ideologies and the resultant regimes. Self-Determination Theory, an empirically based theory that particularly emphasizes the need for autonomy, seems well positioned to support this task.
References


