

Essentialization as a Distinct Form of Abductive Reasoning

Alexios Arvanitis

Business College of Athens

Journal of Theoretical and Philosophical Psychology ®

<http://www.apa.org/pubs/journals/teo/>

This article may not exactly replicate the final version published in the APA journal.

It is not the copy of record.

Author Note

Alexios Arvanitis, Department of Psychology, Business College of Athens.

The author wishes to thank Alexandra Hantzi, Errikos Vassilikos, Antonis Karampatzos, Constantinos Tsinas, Anna Touloumakou and Marietta Papadatou-Pastou for helpful comments on earlier versions of this paper.

Correspondence concerning this article should be addressed to Alexios Arvanitis, Department of Psychology, Business College of Athens, 2 Tatoiou St., 14561, Kifissia, Athens, Greece. E-mail: alexiosarvanitis@gmail.com.

Abstract

Essentialism is often criticized for producing biased behavior. Because it is a view through which people attempt to grasp the essence of things, it appears contradictory that essentialism might result in distortions of reality. Somewhere within essentialist cognitive processes there must be mistakes or omissions that fail to capture reality correctly. In this paper, I treat essentialization as an abductive reasoning process, as a hypothesis, that explains particular characteristics of people on the basis of category membership alone. Besides essentialization, essentialist reasoning can also include deductive and inductive processes that aim at elaborating and testing initial hypotheses. Therefore, essentialist beliefs can be built by essentialization and hypotheses alone or they can be the product of more elaborate reasoning. The relationship of essentialization and general essentialist reasoning to truth and knowledge is discussed.

Keywords: essentialism; abduction; stereotypes; fundamental attribution error; correspondence bias

Essentialization as a Distinct Form of Abductive Reasoning

Essentialism is a philosophical view that dates back to Aristotle's time, when the sophists' relativistic views challenged the very idea of truth. Aristotle wanted to differentiate the true from the seemingly true by focusing on underlying properties (Aristotle, 1955). In the Aristotelian view, the essence reflects the real, *defining* nature of a thing or its kind (Bolton, 1976). In short, the essence makes a thing, or the category it belongs to, what it is. Moreover, essential properties differentiate things from other things (Matthews, 1990) and can therefore be considered a source of establishing commonalities among members of a category and differences among categories. In this sense, not only do essences make category members what they are, but they also establish how they differ from other category members. Essences, therefore, can help us uncover the truth.

On a philosophical level, it is very difficult to find specific essences, that is, specific and defining properties of a person, thing, or category (Cohen, 1978; Matthews, 1990). If philosophers cannot identify specific essences, then people in general may also be unable to do so. However, researchers in the field of psychological essentialism do not study whether or not people are aware of specific essences. They do not examine whether things actually have essences, either. In fact, they treat essences as psychological constructs in order to explain how representations of things are structured: Representations often seem to be structured around "psychological essences", even if they are unknown to the person holding these representations and even if they do not correspond to something that is true and empirical in nature.

I approach essences as psychological properties around which people can construct their representations, without necessarily knowing much about them.

Essences may or may not correspond to something true, but for our purposes this does not matter. I will leave room for the fact that they may correspond to something empirically valid, especially when discussing specific essences, but I will not seek to identify specific defining features that can count as essences. Because essentialist beliefs can exist without knowledge of defining, essential features, I seek to unfold the cognitive processes that result in essentialist beliefs. Indeed, there must be some cognitive processes that consistently result in essentialist beliefs such that it appears people are relying on essences. Unfortunately, there is no consensus on what cognitive processes underlie essentialist reasoning, because researchers in the field have typically emphasized describing and detecting essentialist beliefs (by relying on concepts such as natural kinds or innate potential), rather than pinpointing the cognitive processes that produce them. I focus on *essentialization* as a psychological process and argue that it is an *abductive reasoning process* that explains category member features on the basis of category membership. This theoretical approach opens the door for a re-conceptualization of essentialism that focuses on perception and reasoning. It offers the potential to locate placeholder and specific essentialism along a continuum of knowledge. This may potentially help people avoid the traps of essentialist thinking. It also holds implications for science and the study of cognitive errors, such as the fundamental attribution error (Ross, 1977).

The Search for the Elusive Essence: Specific and Placeholder Essences

Because the essence may not be an explicit element in an individual's cognitive processing, Medin and Ortony (1989) suggest that instead of an essence, we can think of an *essence placeholder* that underlies people's beliefs about a category. Placeholder essentialism is "not the view that *things* have essences but rather the view that people's *representations* of things might reflect such a belief" (Medin & Ortony,

1989, p. 183). Emphasis is often given to representational, causal, placeholder essentialism (Gelman, 2004), which refers to an essence that is causally responsible for the category characteristics and affects their representations without knowledge of the perceiver. This is evident where external and surface properties do not typically determine how something is perceived or even categorized, that is, where external similarity between things is not enough to classify them in the same category. According to a psychological essentialist view, people or things are not perceived to belong in the same category because they are similar; they are perceived to be similar because they belong in the same category (Medin & Ortony, 1989). Similarity is a consequence of something deeper that characterizes the whole category, hence the basis of categorization lies in the deeper, essential features of category members rather than in their perceptual similarity. Ahn, Kim, Lassaline, and Dennis (2000) argue that a *hidden* or unavailable causal feature can be the basis of categorization as long as people infer that such a causally-related feature exists but cannot be explicitly observed.

This is exactly what placeholder essentialism can be treated as: The phenomenon under which a category is perceived and understood on the basis of a hidden underlying causal feature. It does not matter whether people can pinpoint the essence and whether they can understand its causal role in the determination of the category features. In fact, this apparent lack of understanding of essences as causes has led Strevens (2000) to question the need for a concept of essences since *K-laws* alone (K stands for “Kind”), causal relationships that link kind membership to the category’s characteristics without the mediating role of an essence, could stand as explanations for essentialist attributions. There is no need, for example, to speak for the essence of tigerhood to account for a tiger’s stripes in essentialist representations;

Strevens (2000) argues that you can just as easily attribute stripes to simply being a tiger.

Some may contend there are cases when we do have specific essences assigned to categories and people are aware of this. Is not the essence of birds their wings or their ability to fly? Do not some social categories such as men and women have some type of biological basis that makes them who they are? After all, is not water H₂O? In a very constructive contribution, Malt (1994) studied whether people perceive water as H₂O. She found that other factors such as use, location, and source of a liquid play a role in the categorization of the liquid as water. In fact, liquids that were 91% H₂O, such as tea, were not considered water, whereas liquids that consisted of 67% H₂O, such as sewer water, were perceived as water. Consequently, Malt (1994) argues against the strong version of essentialism, where a specific essence accounts for the identity of a category.

It seems difficult to pinpoint a root underlying cause for all observable properties in even the simplest of things, categories, behaviors, or phenomena. Still, there is great difference between identifying a specific essence such as H₂O and not being able to pinpoint any type of essence. By contrast, some essences, like the “DNA”, appear specific enough to account for certain traits but the underlying knowledge--for example, of how specific chromosomes work--is often completely absent, thus leaving room for possible errors. Some essences appear to provide better and more adequate explanations than others. The field of psychological essentialism grapples with developing a theory whose basic concept, the essence, sometimes refers to something concrete and specific, sometimes refers to something hazy and vague, and sometimes is completely absent from the individual's cognition.

Instead of the essence, researchers focus on properties and characteristics of essentialist thinking such as the existence of *inductive* inferences from one category member to another (Gelman & Markman, 1986), the assignment of *innate potential* to category members (Gelman & Wellman, 1991), and an *underlying structure* that is causally responsible for feature characteristics (Ahn et al., 2000). Even social categories can be perceived as if there was an essence underlying the formation of their beliefs, despite the fact that they are social artifacts (Rothbart & Taylor, 1992). One way to approach essentialism for social categories is to examine when people treat social categories as *natural kinds*--categories such as birds, gold, fish--which are found in nature irrespective of human existence and are believed to possess underlying essences that make them different from other categories. In fact, natural kinds' conceptions are quite prevalent in the study of essentialism because they portray a category as inalterable and rich in underlying meaning (Rothbart & Taylor, 1992), as if there was an essence that is so basic to the category members that it does not allow them to change and at the same time offers a wealth of information about them. *Entitativity*, the degree to which a group is seen as a distinct entity, has also been identified as a closely linked, yet distinct, concept to that of essentialism that stresses the perceptual rather than the inferential side of group perception (Demoulin, Leyens, & Yzerbyt, 2006; Yzerbyt, Corneille, & Estrada, 2001; Yzerbyt, Estrada, Corneille, Seron, & Demoulin, 2004; Yzerbyt, Judd, & Corneille, 2004). Both entitativity and natural kinds can be treated as inherent aspects of essentialism (Haslam, Rothschild, & Ernst, 2000). Other aspects such as *universality*, the historical and cross-cultural invariance of categories, have also been identified as integral components of essentialism (Haslam & Levy, 2006).

In their search for the elusive essence, researchers in the field of psychological essentialism focus mostly on whether people develop an understanding of a category as natural, entitative, and invariant. In this sense, they do not really address how people think during the creation of essentialist representations but mostly describe the properties of these representations--inductive potential, innate potential, underlying structure, natural kinds, entitativity, universality--. In this paper, I focus on the actual cognitive process of essentialization, defined herein as a cognitive process that relies on abstract essences in order to explain specific features of category members. I argue that essentialization uses category memberships alone as explanations for the features of category members. I continue to explain the relation of essentialization to more general forms and manifestations of essentialist thinking. More importantly, I discuss the possible errors of essentialist thinking.

The Structure of Essentialist Reasoning

One of the most difficult aspects of studying essentialism is ascertaining what types of cognitive processes are involved. Gelman and Coley (1990) wondered whether children's essentialist inferences involve inductive or deductive reasoning and concluded that they most likely involve inductive reasoning. Roberts (2004) has argued that such inferences are based on abduction, not induction. My approach posits that abduction is the starting point of essentialist reasoning, thereby offering an account that places the search for causality at the center of essentialism. This theoretical approach stresses that essentialist reasoning is initiated by an effort to perceive and explain facts on the basis of *category membership*. Inductive and deductive reasoning, although part of essentialist reasoning, come after this automatic and "instinctive" perceptual process.

In the following paragraphs I will rely on traditional parts of logical arguments, as they were used by Peirce (1992), in order to analyze an example from Gelman and Markman's (1986) seminal work. Gelman and Markman (1986) wanted to examine whether young children would rely on perceptual similarity or category membership in order to predict the behavior of a particular category member. They presented participants with a specific set of cards that showed a picture of a target category member as well as pictures of a same category member that appeared different and of a different category member that appeared alike. On a particular card, participants were shown pictures of a small brown snake, an earthworm, and a cobra. The small brown snake belonged to the same category as the cobra, the category of snakes, but looked like the earthworm, which belonged to the different category of worms. Participants were then asked to predict whether the small brown snake would eat meat--as an earthworm would--or plants--as a cobra would--. Predicting that the small brown snake would eat plants is evidence of essentialist reasoning because such an inference is based on category membership and its underlying properties rather than perceptual similarity. Let us try to examine how a person would decide that the small brown snake would exhibit the behavior of the cobra.

Through essentialist reasoning, a person will conclude that the small brown snake will exhibit the behavior of all snakes, and therefore the same as the cobra, and eat plants. The basic intuitive property of this type of reasoning that I will try to capture by the following analysis is that *there is something about snakes that makes them eat plants*. If there is one thing directly related to the notion of psychological essentialism that should be pinned down, it is the causal relationship between being a snake and eating plants.

Initially, an attribute is assigned to the same category member: The cobra eats plants. From then on, the participant forms a generalized conception that explains why the cobra eats plants. The fact that the cobra eats plants would be explained by the supposition that the behavior of the cobra is a case of a general rule involving snakes: Snakes eat plants in general. This is an example of the formation of a hypothesis, or as Peirce expressed it in his later writings (Peirce, 1998), of abduction. Abduction is the inference of the case from the rule and the result.

Abduction:

Snakes eat plants (Rule)

The cobra eats plants (Result)

∴ The cobra is a snake (Case)

This is arguably how essentialist reasoning is initiated. Through the use of a general rule that connects the feature of the category (eats plants) to the category itself (snakes), individual circumstances (the cobra eats plants) are explained on the basis of category membership (the cobra is a snake)--as cases of a general rule--. According to Peirce, the explanation comes to us by intuition and allows us to explain the result we have witnessed, in this example, the fact that the cobra eats plants.

From then on, essentialist reasoning offers both inductive as well as deductive potential. Deduction is the kind of reasoning children most likely use in order to conclude that the small brown snake eats plants. This form of reasoning establishes the truth of the inference on the basis of the premises being true. It is a valid form of inference, one that necessarily flows from the premises.

Deduction:

Snakes eat plants (Rule)

The small brown snake is a snake (Case)

∴ The small brown snake eats plants (Result)

Although deduction appears valid, it will not provide a sound conclusion if the rule does not correspond to the truth. Induction, on the other hand, is the way in which we test hypotheses about the world. The rule is inferred from particular cases, which lead to a general law. Causality is not part of the reasoning process--as in abduction--but it is indirectly tested by observable data. In other words, induction shows whether a rule is operative.

Induction:

The small brown snake eats plants (Result)

The small brown snake is a snake (Case)

∴ Snakes eat plants (Rule)

Abductive reasoning provides us with intuitive explanations about observed facts. However, the explanation itself cannot be the object of definitive knowledge. It is *hypothetical* and is often called the inference to the best explanation (Harman, 1965). More particularly, though, essentialist reasoning involves causal reasoning processes that use category memberships or category labels as explanations and specific member features as prompts. Essentialist reasoning can thus be manifested as a sub-type of abductive reasoning that provides intuitive explanations for individual results based on category membership. The constructed cause-effect relationship can be used for deductive processes, which appear valid, but are only reliable to the extent that the individual's intuition is accurate. In effect, they are mere predictions, just as the ones Gelman and Markman (1986) asked their participants to make when they asked what a small brown snake would eat. The only way to test hypotheses is through inductive reasoning, which may examine whether the hypothesis can be

supported. In this example, an inductive process would involve observing what the small brown snake would eat.

With regard to the discussion in this section, one may argue the rule “Snakes eat plants” can initially be inferred through induction, not abduction. Although more evidence should be gathered to support the process analyzed is actually abduction, my main argument here is that it should be abduction for it to be defined as essentialist. If it were the mere product of induction, the rule “snakes eat plants” would be a generalization without any underlying cause and would not presume *there is something about snakes that makes them eat plants*. The underlying theoretical cause is only possible if, through abduction, participants explain the result that the cobra eats plants by the rule that snakes eat plants. Otherwise, it is a mere generalization with no explanatory value.

Essentialization and Essentialist Thinking

At this point, I would like to make a distinction between essentialization and general essentialist thinking. Essentialization is the cognitive process that attributes category member features to underlying essences. Since essences are often absent in the individual’s cognition, it can be argued that essentialization is the cognitive process that attributes category member features directly to category membership. I have described this process as a process of abduction that employs category membership for the explanation of a category member feature. The explanation, however, is based on a hypothetical connection between category membership and featured characteristic, not a necessarily acknowledged fact or some sort of real essence. Essentialization is therefore based on an essence placeholder. Essentialist thinking can further use whatever hypotheses are constructed through essentialization, make predictions and test them, resulting either in their acceptance or their rejection.

Essentialist thinking is therefore broader than essentialization and refers mostly to the use of the original category-based hypotheses to make sense of and understand categories and their features.

It should be stressed that not all types of abduction should be associated with essentialization. In fact, not all types of abduction should be connected to essentialist reasoning in general. Only when there is reliance on category membership can we speak of essentialist reasoning, as in the following example:

“White men can’t jump” (Rule)

John is an inadequate basketball player (Result)

∴ John is white (Case)

The formation of any hypothesis that does not rely on category membership is not evidence of essentialist reasoning. For example, the following abductive reasoning is not essentialist because the hypothesis does not rely on category membership, meaning that the case does not simply categorize a person:

Basketball players that do not work hard do not score many points (Rule)

John does not score many points (Result)

∴ John doesn’t work hard at basketball (Case)

Abductive processes that are based on category membership alone are associated with an invisible, elusive essence that hypothetically connects all category members to category features. This particular form of abductive reasoning is what I call essentialization, but essentialist thinking can go a step further: Deduction can be employed to generate predictions and inductive reasoning might serve to test any hypotheses deriving from abduction. If deduction and induction still revolve around the basic hypothesis that links categories to their features, they are elements of

essentialist thinking that goes beyond the initial hypothesis and by doing so, beyond essentialization.

Furthermore, essentialization, just as any type of abduction, is an instinctive response to something surprising, or indeed anything that we observe but cannot explain. In fact, in one of his later writings, this is how Peirce described the abductive process:

The surprising fact, C, is observed;
But if A were true, C would be a matter of course,
Hence, there is reason to suspect that A is true.

(Peirce, 1998, p. 231)

The fact that an explanatory process will be initiated by surprising facts is consistent with recent findings that show that causal reasoning processes are especially triggered when surprising facts are observed, and generally by events that are inconsistent with prior knowledge (Legare, Gelman, & Wellman, 2010). Peirce warns that abduction comes as a flash, as an “act of extremely fallible insight” (Peirce, 1998, p. 227). Its logical structure is equivalent to the structure of “affirming the consequent” (Niinuoloto, 1998) and might therefore be considered a formal fallacy. Abduction is a rather creative form of reasoning that does not conform to strict standards of logic. In fact, “abductive inference shades into perceptual judgment without any sharp line of demarcation between them” (Peirce, 1998, p. 227). Therefore, abduction should not be conceived as a formal reasoning process but as the way in which we perceive surprising facts and automatically form hypotheses in order to explain them. If the causal reasoning process is restricted to abduction, the whole explanatory process is dependent upon an act of insight.

Of the infinite number of hypotheses a person can devise to explain specific circumstances, Peirce was surprised that a person is equipped with the instinctive propensity to make the right guesses (Peirce, 1998). He argued that the capacity to make the right guesses reveals an underlying relation of the person to nature that has developed through evolution. Although there are problems with treating abduction as an instinctual process (Hoffmann, 1999), it may be argued that abduction relies on the evolutionarily-developed ability to grasp the causes of things. Of course the same type of ability could develop through environmental influence early on in development. This kind of ability should be possessed by children, thereby providing an account of how they can make probabilistic assessments about hypotheses (Golpnik & Wellman, 2012) and, especially in the case of essentialization, how they develop an early understanding of essences (Gelman & Wellman, 1991).

Possible fallacies during this process can be avoided by the use of subsequent reasoning. The three reasoning forms described thus far are equivalent to three stages of scientific inquiry: Abduction is equivalent to forming a hypothesis, deduction is equivalent to making a prediction, and induction is equivalent to testing the hypothesis (Flach & Kakas, 2000; Peirce, 1998). Abduction can be thought of as the path from facts toward ideas and induction as the path from ideas to facts that test the ideas (Hoffmann, 1999). If essentialization, then, is followed by the other forms of reasoning, the whole essentialist reasoning process will actually be closer to approaching the truth in a scientific manner and maybe revealing specific essences, should they exist empirically.

The Cognitive “Trap” of Essentialization

A general essentialist reasoning process that utilizes abduction, induction, and deduction tests initial intuitions and builds knowledge. If, on the other hand, an

individual forms a hypothesis and never tests it, but uses it to explain why people or things exhibit characteristics on the basis of group membership, that person will be relying on a rather vague and hidden essence. The possibility that a hypothesis may be rejected through evidence, especially in the realm of science, demands that “a hypothesis adopted by abduction could only be adopted on probation, and must be tested” (Peirce, 1998, p. 95). However, the structure and nature of the essentialist hypothesis often resists testing.

In fact, there is great difference in providing a category-based explanation from providing any other type of explanation. Legare (2012) conducted an interesting experiment in which she examined how children’s types of explanations informed their subsequent exploratory behavior. In particular, participants were asked to explain why a particular light box lit up (or did not). In case of providing explanations that focused on the ways that the light box worked (causal function explanations), children spent more time trying to test their hypotheses by examining the light boxes carefully, opening them or trying out different combinations. However, when providing explanations that simply categorized the light box, as a “blicket” or a “not-blicket”, children engaged in the least exploratory behavior. Consequently, the explanations that employed category membership to account for category member features result in less exploratory behavior, arguably as if categorization itself stands as an explanation. Essentialization in this way can lead to systematic biases.

Essentialist thinking has already been associated with the *illusion of explanatory depth* (Rozenblit & Keil, 2002), the illusion that one has knowledge of underlying theoretical relations and causes when in fact one has no such knowledge. If we view essentialist thinking as a cognitive process that is based on category-based explanations, maybe we can understand how such reasoning withstands further

examination on the basis of manipulative theories of causation, of which most recent prominent example is that of Woodward (2003). These theories suggest that the cause of a phenomenon is that which, if manipulated, would alter/change the phenomenon in question. This is a well-known thesis for experimental psychologists and one that is defended by scientists and non-scientists alike. Of course one would expect that changing the known or unknown essence of something, would definitely alter its features and characteristics. In essentialist representations, category membership is so central to the underlying cause that it comes close enough to this elusive essence. Manipulating category membership is likely to result in altering the phenomenon in question: If the cobra were not a snake, it would probably not eat plants; if John were not white, he would not be bad at basketball. It is true that even if the cobra were not a snake, it might eat plants anyway. And John might be bad at basketball, even if he were not white. All these statements point to the fact that essentialist reasoning may result in false beliefs. Category membership is so central though to the way we understand a person or a thing that in attributing a feature directly to category membership, we create the illusion that we have pinpointed the cause merely due to the fact that any change to category membership would dramatically alter the situation, and therefore the feature or characteristic in question. In terms of the previous examples, category member change is so dramatic that a person would figure that all observable properties like eating plants or being bad at basketball are likely to change as well.

Apart from essentialism offering a seemingly fundamental explanation based on category membership, it is very difficult to actually test this explanation by changing category membership. Imagine how difficult it is to manipulate the chemical structure of certain things, the race, or the gender of people in order to change

category membership. If it were possible, there would probably be an effect on the feature of the category member, but this is rarely tested. The trap of essentialist reasoning lies primarily in the illusion of explanatory depth that category membership as an explanation offers. This type of explanation appears so fundamental as if it were obviously true. Of course John would not be competitive, one might say, if he were not a man and of course Michael would not be depressive if he were not an AIDS patient. No one will attempt to test these hypotheses by changing John's gender or by curing Michael of AIDS. Essentialization poses a significant cognitive trap that offers easy explanations and at the same time makes it very difficult for people to challenge their hypotheses and test them.

Finding Specific Essences

An essentialist hypothesis is not necessarily untrue. However, at the time of its initial formation, the essence is hidden. If people employ further reasoning processes and are able to understand more about the causes of the characteristics of a category, they will not need a hidden essence to explain the features of category members. The hypothesis may be supported by subsequent inductive and deductive processes that aim to uncover the truth. The assumed connection between category membership and observed feature may indeed be validated. In this way, there will be an attempt to replace the essence placeholder--the assumption of a direct connection between category membership and observed feature--with the real and specific essence--the real link between category membership and observed feature--.

For example, an essentialist belief may be that women are sensitive, which is the cause why Jane is sensitive. As such, the belief is based on a hypothetical connection between belonging to the category of women and exhibiting the feature of sensitivity. Deductive processes can generate predictions about Joan and Mary being

sensitive--being women also--and inductive processes can test these predictions. Alternate hypotheses may be constructed as well as different predictions could be made for men through deduction and tested by induction. This process could result in uncovering whether women exhibit the characteristic, and, if they do, what type of connection exists between category and characteristic. The reasoning that aims to uncover the hidden properties of categories and establish the truth of essentialist beliefs constitutes a move from placeholder to specific essentialism. Specific essentialism therefore differs from placeholder essentialism on the basis of correspondence to reality and of knowledge of specific causes of category features.

There is of course great difficulty to assess which essence counts as specific and real enough. DNA, for example, can equally represent a specific or a placeholder essence, depending on how much people actually understand about the role of DNA as the cause of a social category feature. The more people know about the role of DNA, the more they understand what the differences between the social categories are, but also what their commonalities are. With better understanding and extensive knowledge, the category ceases to be a distinct entity that is characterized by an intuitive DNA essence and people either validate their essentialist beliefs by coming to understand a specific and real connection between the development of specific chromosomes and specific category characteristics or abandon the idea of category membership as explanation altogether. Specific essentialism does not refer to naming a specific cause such as "DNA" or "cultural influence", which may even serve as rationalizations for explaining category features away based on category membership; it refers to establishing a *specific connection* between category membership and category feature.

Such specific connections are arguably difficult to establish. To examine differences between men and women, for example, full knowledge extends beyond the simple acknowledgment that there are differences in DNA in order to truly explain the characteristics of men and women. Comprehensive knowledge would include knowledge of all pertinent chromosomes, genes, and hormones as well as social factors including social dynamics, cultural background, economic, political, and even technological environment factors. The more complicated categories and their features are, the more complicated and thorough the use of deduction and induction should be in order to support any category-based hypothesis. Instead, people can simply resort to essentialization, which *does* offer an explanation, although failing to specify the path from category membership to specific member characteristics--in this sense, the explanation might not correspond to the truth--.

There are cases for which essentialist thinking seems appropriate even if relying on essence placeholders, as in the case of natural kinds' categories. It is useful to note that the natural kinds-human kinds distinction may be considered a human construct and, as such, can be treated as a dynamic rather than a static concept (Raskin, 2011). Categories that are based on race or heritability, for example, are very difficult to classify on the basis of this distinction. A rule of thumb is to think of natural kinds as existing outside human discourse. The less something is perceived to be affected by the interaction of humans, the more the category will appear as a natural kind and its properties as deriving from chemical substance. Gold being such a natural kind category, it is definitely easier to explain why it shines rather than why women are sensitive --if indeed they are--. This may also account for why essentialist reasoning is perceived to be more appropriate for use with natural categories than with social categories. It is not so much that the essence of the category has been

thoroughly identified and understood as it is that a real and straightforward cause for the observed characteristics exists. Rothbart & Taylor (1992) argue that people view natural kinds as less arbitrary because of their underlying essences, but maybe people only do so because they can approach the causes of their features intuitively. Even if they do not have accurate knowledge of the causal connection between category membership and characteristic, the assumption that there is a direct connection between category membership and characteristic may be more correct for natural kinds than for human categories. Therefore, an abductive reasoning process that connects the characteristic to the category member may be closer to the truth for natural kinds than for human categories, even if the hypothesis is never followed up.

Human categories on the other hand are very complicated, which makes it difficult to account for the causes of their features by means of abduction alone. This would be equal to doing research in psychology and stopping at the level of the formation of hypotheses. It is often assumed that if people think that DNA causes categories' characteristics, this is evidence of essentialist reasoning. However, recent research attests to the fact that a belief in social determinism can be equally a part of essentialism as a belief in genetic determinism (Rangel & Keller, 2011), meaning that you can essentialize even if you think that social factors shape the essential properties of a category. Essentialization and essentialist thinking in general, therefore, exhibit similar properties whether it is assumed that DNA or some type of societal influence determines the characteristics of a category and its members. It might not really matter what type of connection is assumed between the category and the characteristic if the whole reasoning process primarily stays at the level of an initial abductive reasoning process that explains a particular case in terms of a general rule associating categories with their properties: Such reasoning process should be treated as

essentialization because it relies on an unsubstantiated, intuitive, hypothetical connection between members of categories and their characteristics, the essence placeholder. This type of essence should not be mistaken for the true, specific essence which, if existing, can only be reached by exhaustive examination, especially in human categories.

Human Categories

Essentialism has been linked to stereotype formation and endorsement (Bastian & Haslam, 2006; Haslam, Rothschild, & Ernst, 2002; Yzerbyt et al., 2001) and is generally criticized by social scientists because it is associated with racist and sexist attitudes (Nussbaum, 1993; Sayer, 1997). Categorization based on essences appears to enhance ingroup similarity and between-group differences in ways that can, for example, be used by the majority to reinforce intergroup distinctions and justify discrimination toward minorities (Morton & Postmes, 2009). Mahalingam (2003, 2007) argues that essentialism is used to *naturalize* power relations among different groups (cf. Stoler, 1997). Consequently, essentialism often helps in rationalizing the differential status of social groups. In this way it can be used strategically, for example, for normalizing and privileging specific forms of femininity or for claiming that women have shared characteristics that allow them to act as a collectivity and change the status quo (Stone, 2004).

In fact, when referring to essentialist reasoning, it can be argued that, with regard to social categories, stereotypes are the *rules* that serve to explain a particular *result*. As long as an individual *explains* social phenomena on the basis of stereotypes (cf. Yzerbyt, Rocher, & Schadron, 1997), the reasoning is essentialist. Let us examine Peirce's (1992) initial example of the construction of a hypothesis:

All these beans from this bag are white (Rule)

These beans are white (Result)

∴ These beans are from this bag (Case)

The category “beans from the bag” and its relation to the characteristic “white” are used to explain why *these* beans are white. We can easily think of the rule “All the beans from this bag are white” as an equivalent for a stereotype concerning the category “beans from the bag”. The stereotype is used to explain why *these* beans are white or indeed why any category members might exhibit a particular characteristic. It is also used to predict what type of beans we would expect from this bag, through the following deductive process:

All these beans from this bag are white (Rule)

These beans are from this bag (Case)

∴ These beans are white (Result)

Therefore, based on our initial hypothesis we would expect to pull only white beans out of the bag. If we were not able to do so, then this would mean that our initial hypothesis, as phrased, is not valid. One way to test our hypothesis is to use induction and pull more beans out of the bag. As long as the beans coming out of the bag are white, the hypothesis is supported. If a bean coming out of the bag is not white, different hypotheses can be constructed by the use of abduction.

In order to ascertain the level of accuracy of an essentialist hypothesis, we would need to take all the beans out of the bag. Accordingly, in order to test the truth of a stereotype, we would need to examine all individual category members. This only serves to show that the attainment of true knowledge is very difficult and often impossible. If left untested, beliefs about the causes of characteristics of category members are hypothetical and rest almost exclusively on a reasoning process of abduction: This is a process of essentialization. If an individual does not go further

than these hypotheses or even if that person makes a small step of making predictions based on these hypotheses, the whole reasoning process never really leaves the realm of essentialization. If I hypothesize that George is tall because he is Croatian and I further deduce that any Croatian will be tall, I really have not gone beyond simple essentialization. On the contrary, if I observe other Croatians and examine whether they are indeed tall, I am at least testing the hypothesis through induction. Inductive processes test assumptions about hypothetical rules and are useful in ascertaining the truth of stereotypes or any type of category-characteristic rule--thus having the potential to change them in the light of disconfirming characteristics--. In this way, an individual can come closer to understanding the essences of social categories and validating--or rejecting--essentialist beliefs.

The main example used in this section--the white beans--serves as a metaphor for any type of social category but also shows that essentialism may exhibit the same fundamental reasoning for all types of categories, social or not. It also shows that the same categories may be treated in an essentializing way or a non-essentializing way, depending on whether category membership is used as an explanation of their properties and whether these beliefs are tested.

Some people may essentialize more than others, presumably because of the difficulty of approaching the properties of a category and the simultaneous necessity of providing an immediate and intuitive explanation. It would therefore be expected that having a high need for cognitive closure (Kruglanski & Webster, 1996), that is, a strong desire for definitive knowledge or a high need for cognition (Cacioppo & Petty, 1982), that is, a high need to make sense of the world, would create pressure for an immediate explanation such as the one essentialization offers. Indeed, the above mentioned epistemic needs correlate with beliefs in genetic determinism and social

determinism, which can be viewed as components of psychological essentialism (Keller, 2005; Rangel & Keller, 2011). Furthermore, the need for cognitive closure has been identified as a source of essentialist beliefs (Roets & Van Hiel, 2011), which arguably means that the more we want to have definitive answers about the world around us, the fewer hypotheses we tend to generate (Mayseless & Kruglanski, 1987) and the more likely we are to essentialize and rely only on hypotheses that explain member characteristics on the basis of category membership. At the same time, some categories, such as blacks, women, or AIDS patients are more essentialized than others (Haslam et al., 2000), which would allude to the fact that their properties are complicated and difficult to explain but at the same time, there are situational requirements for providing immediate simplifying explanations that serve as rationalizations for the status quo (cf. Yzerbyt et al., 1997).

The continual use of an abductive reasoning process that explains a particular result as a case of a general rule connecting categories to their properties can offer a general framework of approaching specific categories. Whenever people, for example, explain a woman's behavior by shaking their heads and saying "Women!", they are using essentialist reasoning because they are insinuating that the particular behavior of the woman is adequately explained by (a) the fact that she is a woman, and (b) women behave in ways consistent with the behavior they have just witnessed. Unfortunately these types of remarks are common in everyday life and may be used strategically to promote essentialization and reinforce intergroup distinctions. This type of reasoning distorts the perception of social categories and may produce biased behavior.

Naming the Wrong Causes

Simplifying explanations have been the object of study for researchers in the area of causal attribution, the field which is concerned with how people answer “why” questions and the ways in which they are systematically wrong. One of the most commonly mentioned examples of simplifying explanations is the fundamental attribution error (Ross, 1977), also known as the correspondence bias (Gilbert, 1998), that is, the systematic overvaluing of dispositional explanations over situational explanations. Although there may be a number of reasons why the fundamental attribution error occurs, essentialization might offer a plausible account that focuses on reasoning and its inherent language. Dispositional explanations usually employ the use of adjectives such as “smart”, “handsome”, “introvert” which do not refer to a specific empirical situation and are therefore abstract and decontextualized. Their use supports cognitive economy but resists critical examination, resulting in the fundamental attribution error (Semin & Fiedler, 1988). Such use of adjectives merely categorizes people and can be considered part of an abductive process that aims to offer a hypothetical explanation through category membership. To use an example from Semin and Fielder (1988), Bob can be categorized as dishonest to explain why he is lying. This type of explanation is based on category membership and the general category feature “Dishonest people lie”. Such reasoning is essentialist because it refers to dishonest people *as if* they have a property that causes them to lie and can therefore be used as an explanation for why Bob, a member of the broader category of dishonest people, is lying. In other words, categorizing Bob as dishonest seemingly offers an explanation for his behavior. In this light, essentialization may lead to the fundamental attribution error by focusing on simplifying dispositional explanations, instead of taking the more difficult cognitive path of ascertaining more complicated causes of Bob’s behavior. There are many adjectives, such as “crazy” or “wicked”,

that are often used to explain particular situations by simply categorizing people in a way that seemingly offers an explanation which is not necessarily real and substantiated. This happens very often in the case of mass murderers, whose actions are often explained away by the fact that they are “insane”.

To develop theories, scientists use similar cognitive mechanisms as other people and are prone to the same errors as other people, even children (Gopnik, 1996). Peirce argued that all “why” questions and generally all scientific questions begin with abduction, since it is “the first step of scientific reasoning” (Peirce, 1998, p.106). In order to explain human behavior, such as attacking another person, scientists may also use essentialist thinking and base their explanations on the categorization of the person exhibiting the behavior. Clinical and personality psychology often use category membership to explain behavior: A is schizophrenic, B is an introvert, C is aggressive. In fact, clinicians’ representations of symptoms exhibit causal, theory-like structure instead of simply forming a list of independent symptoms (Ahn & Kim, 2005). Using such trait or mental disorder classification can be viewed as an essentialist explanation if being a member in the categories of schizophrenic, introvert, or aggressive people is seen as the cause of behavior. In supporting these personality explanations there has been a huge “nature versus nurture” debate, which is consistent with the two different types of causes, genetic and social, that essentialist beliefs are founded on (Rangel & Keller, 2011). The fact that the relative contribution of nature and nurture to the majority of behaviors is still under investigation only serves to show that scientists for a long time have supported essentialist explanations without possessing real underlying knowledge. Of course, some scientists see these categories merely as organizing descriptive tools rather than as categories with explanatory potential. But even when categories are used to

construct explanations in an essentialist manner, the explanations themselves might be more correct and better than alternative ones. Still, because essentialist explanations tend to create an illusion of explanatory depth, they should be treated with great caution and tested further than intuition prescribes.

Truth and Knowledge of Essentialist Beliefs

It might be argued that the more we know and understand about the world around us, the less prone we are to essentialization. However, essentialization is not confined solely to an inability to explain characteristics of a category but extends to the reasoning that accompanies it. Our acquired knowledge may never be enough to truly understand social categories, or any categories for that matter. The problem of essentialization arises if we attempt to fill our knowledge gap with intuitive hypotheses alone. The answer to the problem essentialization poses lies not in the depth of our knowledge, but in the ways in which we treat our ignorance.

As long as we refuse to settle for simple hypotheses that connect categories with their characteristics as an explanation of why a member exhibits a characteristic, we avoid basing our judgment on an elusive essence. The best remedy for essentialization is arguably the simple acknowledgement that we do not possess adequate knowledge about categories, especially if we factor in the idea that the way we understand categories and essences is undoubtedly constrained by the way that we construct them (Raskin, 2011). Aristotle, who believed in the essence of things, often produced unfounded hypotheses but at least always added a “perhaps”, as Peirce (1998) points out. Because the need for closure is already considered a source of essentialist beliefs (Roets & Van Hiel, 2011), a refusal to satisfy the desire for definitive knowledge could forestall essentialist reasoning. This is not equivalent to rejecting there is an essence in things. One can believe in the essence of things

without necessarily inferring that a hidden essence is responsible for the observed characteristic of a single member. Instead, one can merely try to understand more and hopefully identify the true and specific essences--should they exist--someday by elaborating on one's knowledge and testing one's hypotheses. The specific essences in this case refer to the true connection between category members and their features. By contrast, a prior acceptance of the existence of a hidden essence and the use of its explanatory power, even when it is performed non-consciously, would cancel the quest for a specific essence by settling for the essence placeholder.

References

- Ahn, W. & Kim, N. (2005). The effect of causal theories on mental disorder diagnosis. In W. Ahn, R. Goldstone, B. Love, A. Markman, & P. Wolff (Eds.) *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin* (pp. 273-288). Washington D.C.: APA. doi:10.1037/11156-015
- Ahn, W., Kim N.S., Lassaline, M.E., & Dennis M.J. (2000) Causal status as a determinant of feature centrality. *Cognitive Psychology*, 41, 361–416. doi:10.1006/cogp.2000.0741
- Aristotle (1955). On philosophical refutations. On Coming-to-be and Passing Away. On the Cosmos (E.S. Forster, D.J. Furley Trans.). Harvard: Harvard University Press.
- Bastian, B., & Haslam, N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology*, 42, 228-235. doi:10.1016/j.jesp.2005.03.003
- Bolton, R. (1976). Essentialism and Semantic Theory in Aristotle: Posterior Analytics, II, 7-10. *The Philosophical Review*, 85, 514-544. doi:10.2307/2184277
- Cacioppo, J.T. & Petty, R.E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116-131. doi:10.1037/0022-3514.42.1.116
- Cohen, S.M. (1978). Essentialism in Aristotle. *The Review of Metaphysics*, 31, 387-405.
- Demoulin, S., Leyens, J. Ph., & Yzerbyt, V. Y. (2006). Lay theories of essentialism. *Group Processes and Intergroup Relations*, 9, 25-42. doi:10.1177/1368430206059856

- Flach, P.A., & Kakas, A.C. (2000) Abductive and inductive reasoning: Background and issues. In P.A. Flach & A.C. Kakas (Eds.) *Abduction and induction: Essays on their relation and integration* (pp. 1-27). Dordrecht: Kluwer Academic Publishers.
- Gelman, S.A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences*, 8, 404-409. doi:10.1016/j.tics.2004.07.001
- Gelman, S.A., & Coley, J.D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology*, 26, 796–804. doi:10.1037//0012-1649.26.5.796
- Gelman, S.A., & Markman, E. (1986). Categories and induction in young children. *Cognition*, 23, 183–209. doi:10.1016/0010-0277(86)90034-X
- Gelman, S.A., & Wellman H.M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition*, 38, 213-244. doi:10.1016/0010-0277(91)90007-Q
- Gilbert, D. T. (1998). Speeding with Ned: A personal view of the correspondence bias. In Darley, J. M.; Cooper, J. *Attribution and social interaction: The legacy of E. E. Jones*. Washington, DC: APA Press.
- Gopnik, A. (1996). The scientist as child. *Philosophy of Science*, 63, 485-514. doi:10.1086/289970
- Gopnik, A., Wellman H.M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin*, 138, 1085-1108. doi:10.1037/a0028044
- Harman, G.H. (1965). Inference to the best explanation. *The Philosophical Review*, 74, 88-95. doi: 10.2307/2183532

- Haslam N., & Levy S. (2006). Essentialist beliefs about homosexuality: Structures and implications for prejudice. *Personality and Social Psychology Bulletin*, 32, 471-485. doi:10.1177/0146167205276516
- Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology*, 39, 113–127. doi:10.1348/014466600164363
- Haslam, N., Rothschild, L., & Ernst, D. (2002). Are essentialist beliefs associated with prejudice? *British Journal of Social Psychology*, 41, 87–100. doi:10.1348/014466602165072
- Hoffmann, M. (1999). Problems with Peirce's concept of abduction. *Foundations of Science*, 4, 271-305. doi:10.1023/A:1009675824079
- Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology*, 88, 686-702. doi:10.1037/0022-3514.88.4.686
- Kruglanski, A.W., & Webster, D.M. (1996). Motivate closing of the mind: "Seizing" and "freezing". *Psychological Review*, 103, 263-283. doi:10.1037//0033-295X.103.2.263
- Legare, C. H. (2012). Exploring explanation: Explaining inconsistent informs exploratory, hypothesis-testing behavior in young children. *Child Development*, 83, 173-185. doi:10.1111/j.1467-8624.2011.01691.x
- Legare, C. H., Gelman, S. A., & Wellman, H. M. (2010). Inconsistency with prior knowledge triggers children's causal explanatory reasoning. *Child Development*, 81, 929-944. doi:10.1111/j.1467-8624.2010.01443.x

- Mahalingam, R. (2003). Essentialism, power and culture: Representations of social class. *Journal of Social Issues*, 59, 733–749. doi: 10.1046/j.0022-4537.2003.00087.x
- Mahalingam, R. (2007). Essentialism, power, and the representation of social categories: A folk sociology perspective. *Human Development*, 50, 300-319. doi:10.1159/000109832
- Malt, B.C. (1994). Water is not H₂O. *Cognitive Psychology*, 27, 41-70. doi:10.1006/cogp.1994.1011
- Matthews, G.B. (1990). Aristotelian essentialism. *Philosophy and Phenomenological Research*, 50, 251-262. doi:10.2307/2108042
- Mayseless, O., & Kruglanski, A. W. (1987). What makes you so sure? Effects of epistemic motivations on judgmental confidence. *Organizational Behavior and Human Decision Processes*, 39, 162-183. doi:10.1016/0749-5978(87)90036-7
- Medin, D.L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). Cambridge: Cambridge University Press.
- Morton, T. A., Hornsey M. J., & Postmes T. (2009). Shifting ground: The variable use of essentialism in contexts of inclusion and exclusion. *British Journal of Social Psychology*, 48, 35-59. doi:10.1348/014466607X270287
- Niinuloto, I. (1998). Defending abduction. *Philosophy of Science*, 66, S436-S451. doi:10.1086/392744
- Nussbaum, M.C. (1993). Social justice and universalism: In defense of an Aristotelian account of human functioning. *Modern Philology*, 90, S46-S73. doi:10.1086/392123

- Peirce, C. S. (1992). *The Essential Peirce: Selected Philosophical Writings, vol. 1 (1867-1893)*. Bloomington and Indianapolis: Indiana University Press.
- Peirce, C. S. (1998). *The Essential Peirce: Selected Philosophical Writings, vol. 2 (1893-1913)*. Bloomington and Indianapolis: Indiana University Press.
- Rangel, U., & Keller, J. (2011). Essentialism goes social: Belief in social determinism as a component of psychological essentialism. *Journal of Personality and Social Psychology, 100*, 1056-1078. doi:10.1037/a0022401
- Raskin, J. (2011). On essences in constructivist psychology. *Journal of Theoretical and Philosophical Psychology, 31*, 221-239. doi:10.1037/a0025006
- Roberts, L.D. (2004). The relationship of children's early word acquisition to abduction. *Foundation of Science, 9*, 307-320. doi:10.1023/B:FODA.0000042845.38424.a3
- Roets, A., & Van Hiel, A. (2011). The role of need for closure in essentialist entitativity beliefs and prejudice: An epistemic needs approach to racial categorization. *British Journal of Social Psychology, 50*, 52-73. doi:10.1348/014466610X491567
- Rozenblit L.R., & Keil F.C. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science, 26*, 521-562. doi:10.1016/S0364-0213(02)00078-2
- Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 173-220). New York: Academic Press.
- Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kinds? In G.R. Semin & K. Fiedler (Eds.), *Language and social cognition* (pp. 11-36). London: Sage.

- Sayer, A. (1997). Essentialism, social constructionism and beyond. *The Sociological Review*, *45*, 453-487. doi:10.1111/1467-954X.00073
- Semin, G. R., & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology*, *54*, 558-568. doi:10.1037//0022-3514.54.4.558
- Stoler, A. L. (1997). On political and psychological essentialisms. *Ethos*, *25*, 101–106. doi:10.1525/eth.1997.25.1.101
- Stone, A. (2004). Essentialism and anti-essentialism in feminist philosophy. *Journal of Moral Philosophy*, *1*, 135-153. doi:10.1177/174046810400100202
- Stevens, M. (2000). The essentialist aspect of naïve theories. *Cognition*, *74*, 149-175. doi:10.1016/S0010-0277(99)00071-2
- Woodward, J. (2003). *Making things happen: A theory of causal explanation*. New York: Oxford University Press.
- Yzerbyt, V., Corneille, O., & Estrada, C. (2001). The interplay of subjective essentialism and entitativity in the formation of stereotypes. *Personality and Social Psychology Review*, *5*, 141-155. doi:10.1207/S15327957PSPR0502_5
- Yzerbyt, V. Y., Estrada, C., Corneille, O., Seron, E., & Demoulin, S. (2004). Subjective essentialism in action: Self-anchoring and social control as consequences of fundamental social divides. In V. Y. Yzerbyt, C. M., Judd, & O. Corneille (Eds.), *The psychology of group perception: Perceived variability, entitativity, and essentialism* (pp.101-126). London: Psychology Press.
- Yzerbyt, V. Y., Judd, C. M., & Corneille, O. (2004). Perceived variability, entitativity, and essentialism: Introduction and overview. In V. Y. Yzerbyt,

C. M., Judd, & O. Corneille (Eds.), *The psychology of group perception : Perceived variability, entitativity, and essentialism* (pp.1-24). London: Psychology Press.

Yzerbyt, V., Rocher, S., & Schadron, G. (1997). Stereotypes as explanations: A subjective essentialist view of group perception. In R. Spears, Oakes, P.J., Ellemers, N., & S.A. Haslam (Eds.), *The Social Psychology of stereotyping and group life* (pp. 20-50). Cambridge: Blackwell.