

**The Original Sin of Cognition:
Race, Prejudice and Generalization**
Sarah-Jane Leslie

To appear in *The Journal of Philosophy*

Draft: Please do not cite or circulate

Introduction

Long before we learn to talk, our expectations concerning novel members of a category are shaped by our experience with other members. We expect, for example, that objects that share obvious perceptual qualities will share their more hidden properties too. If the one rattles when shaken, we expect the other will too (Baldwin, Markman, and Melartin 1993). By our first birthday, these inductive inferences are guided by language; we expect that even superficially dissimilar objects will share their hidden properties if they are identified by the same common noun; if, for example, each is introduced as ‘a blickett’ (Graham, Kilbreath, and Welder 2001). From the very beginning, we are disposed to generalize from experience with a given item to other items that we perceive as belonging to a common category.

There is, presumably, some innate cognitive mechanism that is responsible for these early inductive generalizations. I argue here and elsewhere that generics – sentences such as ‘ravens are black’ and ‘tigers are striped’ – express the generalizations that are delivered by this basic mechanism. If this is so, then generics provide us with a window onto the workings of this mechanism. This paper is concerned with a particular aspect of this mechanism, namely the manner in which general conclusions regarding dangerous or harmful information are reached. I will argue that the workings of this mechanism have played a significant role in the formation of negative racial stereotypes, and will survey a

variety of historical sources that support this claim. This paper in no way purports to be an exhaustive analysis of all the factors that have formed and sustained racist attitudes over time. Rather, I identify and discuss *one* particular cognitive bias that has given birth to many a prejudice.

Default Generalizations

Generics, for example ‘birds lay eggs’, ‘tigers are striped’, and ‘cats have whiskers’, express generalizations. They are not about particular individuals, but rather say something about the entire category mentioned in them. The judgments they express involve commitments as to the properties of previously unencountered instances of the kind. Generics are not the only linguistic devices we have for making general claims, and correspondingly, generic generalizations are not the only kind of generalization we can make or understand. For example, the quantifiers ‘all’ and ‘every’ express universal, exceptionless generalizations. The quantifier ‘most’ is used to make generalizations concerning more than half of the instances of the category in question, while ‘some’ concerns but a few instances. We may express statistical generalizations using operators such as ‘less than forty-two percent of’, and so on. We have many ways of speaking and thinking about the general rather than the particular.

Of all these means of expression, generics have proved especially puzzling for linguists and philosophers of language. Their truth conditions seem to be enormously complex. Why, for example, is ‘birds lay eggs’ true, while ‘birds are female’ is false? It is, after all, only the female birds that lay eggs. And why is ‘mosquitoes carry the West Nile virus’ true, and ‘books are paperbacks’ false given that less than one percent of mosquitoes carry the virus while over eighty percent of books are paperbacks? Such

puzzling examples abound. Theoreticians have employed everything from non-monotonic logics to iterated possible world modalities and comparative probability measures in an attempt to understand these sentences. Each and every one of these accounts is subject to counterexample (Leslie 2007, 2008).

In contrast to the theoretician's woes, the young child finds generics quite easy to master. Generics appear in children's speech as early as age two, long before children regularly use explicit quantifiers such as 'all' and 'every.' While any linguist can quickly provide a concise definition of 'all' in terms of the subset relation, and cannot define generics even with all the machinery of possible worlds and non-standard logics, children find the elusive generic far more comprehensible than this seemingly transparent 'all.' At age three, not only are children producing generics frequently, and explicit quantifiers rarely, they even sometimes interpret explicit quantifiers as though they were generics. In a recent study, Susan Gelman and her colleagues found that three-year-olds are already at adult levels of competence in their understanding of generics. Even more surprising, though, is that the three-year-olds interpreted explicitly quantified statements as though they were generics! Finding the explicit quantifiers difficult to understand, they responded to the experimenters as though confronted with generics in all cases (Hollander, Gelman, and Star 2002).

It is even more puzzling that children acquire generics so easily when we notice that there is not even an articulated generic operator. We do not say 'Gen tigers are striped,' in the way that we would say 'most tigers are striped.' To express a generic, we simply omit a quantifier. To acquire generics, it would seem, children must not only isolate and learn some intractably complex set of truth conditions, they must also learn to

associate this set with the *absence* of a quantifier. (Associations with absence are, of course, notoriously difficult to master.) This is no quirk of English either; there is no known language that has an articulated and dedicated generic operator.

In Leslie (2007, 2008), I argue that this ‘paradox of generic acquisition’ is resolved by understanding generics to express our conceptual system’s default mode of generalization. The ability to generalize predates the acquisition of language, so there must be some mechanism for generalization already present by the time the child begins to speak. Generics simply give voice to these most basic and cognitively primitive generalizations. I argue that the strange and puzzling truth conditions of generics are traceable to quirks and biases in this mechanism. The child, already in possession of the mechanism, does not need to identify and learn any complicated truth conditions. She need only continue generalizing by way of that basic mechanism she has been using since her earliest days.

This view also explains the otherwise puzzling linguistic universal: why is it that no language has a generic operator? I suggest that this basic mechanism constitutes the mind’s default means of generalizing, and the cognitive system must be *told*, as it were, *not* to use it. An explicit use of, say, ‘all’ informs the mind that it is to deviate from its default, and form a *universal* generalization instead of a default one. Absent such an explicit instruction, the cognitive system relies on its default. The unmarked generic allows the mind’s default mode of operation to proceed unchecked.

If I am correct, and generics express the cognitive system’s default mode of generalizing, then the study of generics may provide us with some insight into the workings of the mind. Elsewhere I have used facts about cognition to shed light on the

semantic behavior of generics, but it may be fruitful to also work in the other direction, using generics as a window onto the mind's unconscious operations. The ability to generalize is a fundamental cognitive ability, so understanding our default mode of generalizing has the potential to illuminate much of our thinking. We might expect that, once we have isolated characteristics of this default mode, we will see these characteristics recurring throughout our reasoning. Thus, while I will use generic sentences to identify some characteristics of our default mode of generalizing, this paper will not be concerned with language, but rather with more general patterns of thought.

Striking Property Generics

In Leslie (2008), I identified various classes of generics, one of which I termed 'striking property generics'. This class includes sentences such as:

Mosquitoes carry the West Nile virus
Sharks attack bathers
Deer ticks carry Lime Disease
Pit-bulls maul children
Tigers eat people

These sentences are intuitively true, even though very few members of the kind in question possess the predicated property. As it happens, less than one percent of mosquitoes carry the West Nile virus, and yet we are quick to assent to 'mosquitoes carry the West Nile virus,' even after learning this statistical fact. (Conversely, 'mosquitoes don't carry the West Nile virus' remains patently false, even though 99% of mosquitoes don't carry the virus.)

It appears that these generics require for their truth only that *some* of the kind possess the property in question.¹ This is not true for generics in general; for example,

¹ One might be tempted to think that these generics are true because, e.g., *only* mosquitoes carry the West Nile virus. There is certainly a reading of 'mosquitoes carry the West Nile virus' to that effect; to see this,

some cats are female, but ‘cats are female’ is false, and *some* (in fact, *most*) mosquitoes don’t carry the West Nile virus, but the corresponding generic is false. Such examples abound. I suggest that the generics above are special in that their predicates express properties that we have a vested interest in avoiding. If even just some small portion of a kind possesses a property that is harmful or dangerous, then a generic that attributes that property to the kind is likely to be judged true.

Since we are working under the hypothesis that generics express our default mode of generalizing, this observation implies that our most primitive way of dealing with dangerous or harmful information involves the rapid generalization of this information to the appropriate category. We do not wait around to see how many tigers in fact eat people, or what percentage of tigers eat people before drawing a general conclusion – even a single instance may be enough for us to conclude that tigers eat people. It is not hard to see the evolutionary benefits of such a disposition, since the costs of under-generalizing such information are potentially huge. Our ancestors were far better off jumping to conclusions, as it were, rather than sticking around to judiciously determine the precise likelihood of their being eaten.

The tendency to rapidly generalize such striking information manifests itself elsewhere in our thinking. Consider, for example, how many murders one must commit to be a murderer, versus how many times one must worry to be a worrier. The latter case

try stressing ‘mosquitoes’, as in ‘MOSQUITOES carry the West Nile virus,’ or else paraphrasing the sentence as ‘it is MOSQUITOES that carry the West Nile virus.’ Consider, however, whether one’s intuitions would change upon learning that deer ticks also carry the virus. This would falsify one reading of the sentence; an assertion of ‘it is MOSQUITOES that carry the West Nile virus’ can be countered by the observation that deer ticks do likewise. However, there is still a salient reading of the sentence upon which it remains true. There is nothing contradictory, or even infelicitous, about the remark that mosquitoes carry the West Nile virus, and deer ticks do too. Thus these generics do not depend on the property’s being uniquely possessed by the subject.

requires one to worry with considerable regularity, whereas a single murder suffices to make one a murderer.

Our perception of others is also mediated by this tendency. A single very evil act in the course of a person's otherwise respectable life is enough for us to count the person as 'bad.' People who learn that a neighbor or colleague has committed a murder, are often heard to remark that the perpetrator '*seemed* like such a nice person.' The implication is, of course, that he only appeared to be a nice person; committing a murder being, of course, incompatible with actually being a nice person. The response is never to the effect of 'what's one little murder compared with a lifetime of decency?' Attempts to mitigate the perception of the perpetrator as morally bankrupt by pointing out his virtues ('he is a good father, and kind to dogs') are seen as either pathetic, or in extremely bad taste. One significant act of evil can obscure many good deeds a person has performed. We do not waste time comparing the regularities with which he performed good deeds as opposed to bad.

The disposition to generalize strikingly negative information on the basis of even a single event thus appears to be a pervasive aspect of our thinking. For generalizations concerning neutral or positive information, we require the instances or events to occur with a significant regularity; this is not so with negative information. There is a fundamental asymmetry between the impact of very negative information and the impact of neutral or more positive information on our intuitive generalizations.²

² The same arguably applies to strikingly positive information. One extremely large charitable donation presumably suffices to make one a philanthropist, though if it is a one time occurrence, the donation must be very large indeed. (A single gift of a moderate sum does not a philanthropist make, unfortunately.) Such cases are less clean-cut, however, and examples are far less readily available.

Complications and Refinements

Let us consider striking property generics in more detail. Their truth conditions are not quite as straightforward as the foregoing discussion suggests. We have been speaking as if a generic ‘Ks are F’ is true iff some Ks are F, given that *being F* is a dangerous or harmful property. But this would suggest that ‘insects carry the West Nile virus’, or even ‘animals carry the West Nile virus’ would also be true – certainly there are some insects, and therefore some animals that carry the virus, namely those few unfortunate mosquitoes. Similarly, the truth of ‘tigers eat people’ would entail the truth of ‘mammals eat people’, and from the truth of ‘sharks attack bathers’ we should conclude that sea creatures attack bathers. People do not tend to find these inferences acceptable, so the truth conditions of these generics cannot be as simple as our original analysis indicates.

In Leslie (2008), I argue that the mechanism of generalization in question seeks a *good predictor* of the property in question. It is easy enough to see an evolutionary rationale behind generalizing striking properties only so far up the taxonomic hierarchy. If our ancestors had undertaken to avoid all mammals after seeing a tiger eating one of their companions, the costs of doing so may well have outweighed the benefits. (One could waste a lot of time running from small harmless creatures.) Someone who avoided all animals, big or small, after witnessing a lion maul his companion would be at a significant disadvantage relative to a more sophisticated competitor who limited his conclusions to lions alone.

An efficient generalizing mechanism, we might suppose, should seek a good predictor of the striking property – a kind that is inclusive enough to aid us in avoiding the property, but not so inclusive as to needlessly hamper our activities.

I further suggest that what makes a kind a good predictor of a striking property is that the members of the kind that do not possess the property are typically *disposed* to possess it (Leslie 2008).³ It matters, then, for the truth of ‘mosquitoes carry the West Nile Virus’ that the virus-free mosquitoes will carry the virus if circumstances allow. ‘Sharks attack bathers’ is true only if the sharks that never in fact cause harm to humans would typically do so given half a chance, and so on. Statements such as ‘animals carry the West Nile Virus’ and ‘sea creatures attack bathers’ are false, because the members of the kinds in question do not share the relevant dispositions.

A generic statement in which a striking property is predicated is, I claim, true if and only if some members of the kind in question possess the relevant property, and the others are typically disposed to possess it. If this is correct, then we must allow for the possibility that some of the striking property generics listed above are, in fact, strictly false. Perhaps it is only Great White Sharks that are disposed to attack bathers (as it is sometimes claimed), or perhaps only mosquitoes with a particular mutation are capable of carrying the virus. If these turn out to be the facts, then my account predicts that the above generics are in fact false, and it is only the weaker claims ‘Great White Sharks attack bathers’ and ‘mosquitoes with a particular mutation carry the West Nile Virus’ that are true. This seems to me the intuitively correct conclusion here: under such circumstances, the more inclusive generic claims are, strictly speaking, false.

To determine which striking property generics are strictly true and strictly false, then, would require some rather detailed knowledge of dispositions and capacities. We do not normally possess such knowledge, and so, as always, we do the best we can with

³ The metaphysics of dispositions is an intriguing subject, and I shall not delve into it here, but rather rely on our intuitive understanding of the notion. For two recent discussions of dispositions, see Fara (2005) and Lewis (1997).

what we have. To probe this further, let us set aside questions of whether sentences are true or false, and consider how our basic mechanism of generalization must work, if my theses here are correct.

Dispositions, Essences, and Basic-Level Kinds

If the preceding remarks are correct, then our basic mechanism of generalization, when confronted with the manifestation of a striking property, seeks to generalize that property to a kind whose members are disposed to manifest it. Detailed scientific knowledge of dispositions is not likely to be available in most instances of these generalizations. There is, however, much evidence to suggest that our conceptual framework provides loci within our psychological taxonomy of kinds at which we expect there to be a high degree of shared natures and dispositions. These kinds – known as *basic-level kinds* – are, I suggest, the default subjects of our striking property generalizations. Absent information to the contrary, we assume that the basic-level kind is a good predictor of the striking property. We assume that its typical members share the relevant disposition, unless we learn otherwise.

The notion of a basic-level kind is an important explanatory notion in cognitive psychology. The term was coined by Rosch (1978), who found that various measures of psychological salience converged on a particular taxonomy that is psychologically privileged. For example, when asked to identify what is in a picture of the animal that is Princeton's mascot, people tend to identify it as a tiger, rather than as a Bengal tiger, or as a mammal, or a vertebrate, and so on. Alternatively, if people are asked to "list features" that they associate with various kinds, basic-level kinds have the greatest number of

features that are both widely shared by members of the kind and *not* shared by members of comparable kinds.

For kinds that are taxonomically below the basic level (so-called subordinate kinds), the features listed tend to be ones that are listed for other subordinate kinds; Bengal tigers share most of their psychologically salient features with other types of tigers. If asked to list features for superordinate kinds such as *mammal*, people have much greater difficulty coming up with features, and often list features that are not widely shared by members of the kind.

The notion of a basic-level kind is an explanatory psychological notion, not an explanatory biological notion. In many cases, the basic-level kind corresponds to genus or a species considered as biological taxa, but this is not always the case. For example, while the basic-level kind tiger corresponds to a species of the genus *Panthera*, the basic-level kind jellyfish corresponds to the class *Scyphozoa*, which has many orders, families, genera, and species below it.

In the course of language acquisition, names for basic-level kinds are learned first, and there is a high degree of cross-cultural agreement about basic-level taxonomy, even though cultures may differ significantly on the taxonomy of superordinate and subordinate kinds. The most interesting feature of basic-level kinds from our point of view, however, is that they are seen as having rich inductive potential. Children as young as two readily generalize information across members of a basic-level kind, even if the members are perceptually dissimilar (Gelman 2003). Basic-level kinds thus appear to be privileged starting points for our common inductive generalizations.

Susan Gelman and her colleagues have argued that our folk theories of biology treat the folk biological kinds as *essentialized*, particularly if they are basic-level kinds (Gelman 2003; Medin and Ortony 1989). To view a kind as essentialized is to form the (tacit) belief that there is some hidden, non-obvious, and persistent property shared by members of that kind that causally grounds their common properties and dispositions. One might believe, implicitly or explicitly, that there is something about tigers that causes them to have stripes, to growl, to hunt their prey, and so on. These are not accidental features of tigers; they are grounded in the very nature of tigerhood. What is more, we believe that even a crippled albino tiger possesses this intrinsic, “essential” nature, even if she does not manifest its outward effects. The essence of tigers causally grounds these dispositions, though does not guarantee their manifestation, since adventitious factors may intervene. (The relevant notion of essence at work in the psychological literature is obviously not the philosopher’s stricter notion of that intrinsic aspect of a thing which grounds all and only the intrinsic metaphysical necessities that hold of the thing.)

The notion of viewing a kind as essentialized may be a little vague, but is intuitive nonetheless. It should be noted that the claim being made here is that our folk theories treat biological kinds in this way, not that this is the correct metaphysics of these kinds. The claim is merely psychological.

To the extent that a kind is essentialized, we are disposed to treat it as supportive of inductive generalizations and inferences about non-obvious properties. In this respect, basic-level kinds are the most essentialized, in comparison with both subordinate and superordinate kinds. Children as young as three expect that members of the same basic-

level kind will have the same internal organs, even if some of the members look quite different from the others (Gelman 2003). Young children have equally strong views about the power of nature over nurture when it comes to cross-species comparisons; they expect that a cow raised from birth by pigs will look like a cow, say moo, and so on (Gelman and Wellman 1991). They also maintain that a raccoon dressed up a skunk is nonetheless a raccoon, and possessed of raccoon innards, thus demonstrating a belief that there is more to kind-membership than outward appearance (Gelman and Wellman 1991). Gelman and her colleagues argue at length that these convictions reflect a belief in the essences of these kinds. From a very young age, we think that there is something intrinsic to the nature of cows that will cause it to resemble other cows, regardless of how it is raised.

I suggest that, in making a striking property generalization, we choose by default the highly essentialized basic-level kind as the locus of the generalization. These kinds are the primary loci of our inductive generalizations and inferences, so it is perhaps not surprising that they are also the loci of these particular generalizations. Basic-level kinds are fine-grained enough to reduce wasted effort, but still coarse-grained enough to allow one to err on the side of caution. And since they are the most readily recognizable categories, they make for a very practical starting point for striking property generalizations.

Most importantly, however, basic-level kinds – because they are *highly* essentialized – are supportive of inferences concerning the dispositions of its members. Upon seeing a tiger eat a companion of ours, we conclude that there is something about tigers that disposes them to eat us – it lies in their nature to eat us, given half a chance. A

typical tiger is thereby, i.e. thanks to his underlying intrinsic nature, disposed to eat us. (If one is unconvinced by this example, consider the following. A Savannah cat is a hybrid bred from the domestic cat and an African hunting cat known as the serval. The adult Savannah is quite magnificent, weighing in at 40lbs, and possessed of a long neck and beautiful leopard-esque spots. There have been no documented cases (to my knowledge) of Savannahs attacking either people or other pets, though the breed has only been around for a few years, so its general tendencies are not yet well understood. Suppose, however, that tomorrow we see splashed across the New York Times a report of a Savannah savagely attacking a toddler. Would one suspend judgment as to the typical Savannah's violent tendencies, or would one immediately view the entire breed as dangerous?)

When an instance of a particular basic-level kind manifests a striking property, it seems that by default we take the manifestation of that property to be grounded in some nature common to the members of that basic-level kind. Unless we learn otherwise, we therefore take the disposition to manifest the property to belong to the typical members of the kind, the one's that share the common nature, and are free of any adventitious inhibitor of the property's manifestation.

It must be noted that non-striking properties are not treated in this way. Upon learning that a given tiger is female, we do not conclude that typical tigers, by nature, are disposed to be female – a disposition that simply fails to manifest itself in the case of male tigers. If we saw a hamster with an odd growth on its back, it would not occur to us to decide that the disposition to grow such lumps is grounded in hamsterhood, though rarely manifested. The rapid generalizations to typical underlying dispositions are specific to striking properties.

Race and Essence

In recent years, a number of social psychologists have argued that we view racial kinds as essentialized in much the way we view basic-level animal kinds. This line of thinking was popularized by Rothbart and Taylor (1992), though the observation originated with Gordon Allport in 1954, who wrote that:

... a belief in essence develops. There is an inherent 'Jewishness' in every Jew. The 'soul of the Oriental,' 'Negro blood,' ... 'the passionate Latin' – all represent a belief in essence. A mysterious mana (for good or ill) resides in a group, all of its members partaking thereof (Allport, 1954: 173-174, reprinted in Rothbart and Taylor 1992).

Rothbart and Taylor argue that we often act as though we view racial kinds as essentialized biological kinds. There is also growing evidence that young children essentialize racial kinds in ways that parallel their essentialization of basic-level animal kinds. For example, young children understand that race is heritable, and it dominates their thinking about other heritable characteristics, such as body build, if they are asked to decide which depicted children belong to which depicted parents (Hirschfeld 1996). As adults, we readily accept that skin color is at best a defeasible guide to race; we accept that a person may *really* be black, even if he looks white. The 'one-drop' rule in the United States for determining racial identity would strike us as utterly ludicrous – perhaps even unintelligible – were we not disposed to conceptualize race this way. Just as with natural kinds, we treat race as inhering in more than just superficial, readily observable characteristics, allowing that there may be some deeper fact of the matter waiting to be discovered.

The discovery that, from the point of view of genetics, race is not a useful biological category, has been poorly assimilated into people's thinking. Most people find

the idea stunningly counterintuitive, to the point of being almost obviously false. Even amongst researchers, the project of seeing, e.g., whether there are differences in IQs between racial groups persists to this day, despite the overwhelming evidence that racial categories do not carve nature at its genetic joints and are therefore not explanatory difference-makers for something like IQ. A belief in the essentiality of racial categories persists.

While the observation is often made that we see racial kinds as essentialized, rarely is any explanation offered of why this should lead to racism. Discussions often proceed as though it was self-evident why essentialized beliefs should lead to negative racial stereotypes and prejudice. *Essentialization alone does not in any way guarantee prejudice.* We essentialize animal kinds such as cats and dogs without any tendency to bear them ill-will, and in various periods of history people have essentialized their own groups as a means of self-glorification (the ancient Greeks were perhaps such an example, and the Nazis' view of the Aryan race is undoubtedly such an example). Thus essentialization of a kind does not inevitably lead to prejudicial bias. There is no intrinsic, necessary connection between racialism – the belief that there are essential differences between the races – and racism. What aspects of our thinking form the bridge between racialism and racism?

The aim of this paper is to develop a novel explanatory hypothesis concerning the transition from racialism to racism.

Fear, Race, and Generalization

The cognitive disposition to rapidly generalize strikingly negative information helps bridge the explanatory gap between race essentialism and racism, or so I will argue. I

should reiterate that what follows is in no way intended as an exhaustive account of the cognitive factors underlying racism, and it barely touches on the social and cultural factors that have bred and sustained racism. Rather, I seek to identify a particular and pervasive aspect of our thinking, itself having no intrinsic connection to either racialism or racism, which nonetheless leads us down the dark road to prejudice. Our most primitive method of generalization, when coupled with an essentialized view of racial kinds, gives rise to a primitive form of racist thinking.

The basic idea is quite simple: just as it takes but a few instances of sharks attacking bathers, or of mosquitoes carrying the West Nile virus, for us to make the corresponding category-wide generalization, so also a strikingly negative action on behalf of a few members of a racial minority may lead us to a general belief concerning their entire racial category. A belief in the essentialized nature of racial categories makes them appear to be appropriate loci of striking property generalizations; our disposition to thoughtlessly generalize dangerous or harmful information does the rest.

If racial kinds are indeed seen as sufficiently essentialized to be the locus of striking property generalizations, we have the disturbing consequence that, as a result of this profound and pervasive cognitive bias, a few appalling acts on behalf of some members of a given race can lead us to draw conclusions about the race in general. This general belief applies to all members of the race; from it we can conclude that any given typical member is disposed to commit such acts in virtue of the nature they share with their cohorts. Extreme and aberrant actions on behalf of the few can thus lead to conclusions concerning all.

This pattern of reasoning is elegantly illustrated in Laura Ingalls Wilder's *Little House on the Prairie*. The Ingalls' rather prejudiced neighbor, Mrs. Scott claims that "The only good Indian is a dead Indian" – certainly a very sweeping and inclusive claim. She immediately justifies the claim by citing the Minnesota Massacre, going as far as to say "To anyone who disagrees, I say, 'remember the Minnesota Massacre!'" Mrs. Scott reasons from the single horrific incident of the Minnesota Massacre to the conclusion that there are no good (living) Indians. Considering that the Scotts and the Ingalls were living in Kansas at the time, it is unlikely that she believed any of the Indians they actually encountered had any personal involvement in the Minnesota Massacre. Nonetheless, she took the single incident to justify the universal claim that the only good Indians were dead Indians. She was also apparently confident that only someone who had forgotten about the massacre would disagree with her on the point.

Contemporary and Historical Examples

The September 11 Attacks

Mrs. Scott's reasoning is a perfect illustration of how human beings can move from a horrific particular to a sweepingly prejudiced generalization. If reasoning of this sort really is a pervasive cognitive disposition, evolved over the ages, then we should find many examples of it, both today and throughout human history. We should not be surprised if these generalizations hovered perpetually in the background whenever mankind was formulating prejudiced attitudes towards essentialized kinds. Sadly enough, we do not need to look to the distant past to find examples.

Nothing has done more to harm the plight of Muslims in America than 9/11. Of course *Muslim* is not a racial category, but just as racial categories present as basic-level

kinds so do ethnic and religious categories, and the spectrum of primitive prejudice thus includes ethnic and religious prejudice as well as race prejudice. Religious kinds have long been treated on a par with racial ones. (Consider, for example, historical and contemporary attitudes towards Jews.) It is also not clear how well most people hold in mind the distinction between Muslims and Arabs, or Muslims and Middle-Easterners more generally. The distinction may be sufficiently blurred to allow people to transfer racial modes of thinking to this particular religious kind.

In the aftermath of 9/11, hate crimes against Muslims rose more than 1,600%, according to FBI statistics (reported in the San Francisco Chronicle, <http://sfgate.com/cgi-bin/article.cgi?f=/c/a/2002/11/26/MN224441.DTL>). Hate crimes are, by definition, crimes motivated by the mere fact that the victim is a member of a particular group; the hate crimes following 9/11 were motivated by the fact that the victims in question were Muslims. Many of these crimes were committed against Muslim women and children; the perpetrators surely were not under the impression that their victims were themselves personally responsible or even involved in the 9/11 bombings. It was sufficient that the victims were Muslims. (There were, however, a good few instances in which the victims were Sikhs, rather than Muslims. Reports of these instances were often greeted with eye-rolling on the part of the cognoscenti: ‘what ignorance to target the innocent Sikhs, who had nothing to do with 9/11!’ This is supposed to *contrast* with the crimes that successfully targeted uninvolved Muslims?) We might characterize the reasoning of the hate crime perpetrators as moving from the horrific events of 9/11 – events which involved a rather small number of extreme

individuals⁴ – to the conclusion that the arbitrary Muslim deserved to be victimized in virtue of being Muslim. The conclusions drawn from the 9/11 attacks did not concern just the bombers and their supporters, but concerned Muslims in general.

Such generalizations were made even by members of Congress. Shortly after 9/11, Representative John Cooksey told a Louisiana radio station, “If I see someone [who] comes in that's got a diaper on his head and a fan belt wrapped around the diaper on his head, that guy needs to be pulled over.” In Georgia, Representative C. Saxby Chambliss told law enforcement official to “just turn [the sheriff] loose and have him arrest every Muslim that crosses the state line” (reported by the Human Rights Watch, <http://www.hrw.org/reports/2002/usahate/usa1102-05.htm>). These statements again reflect conclusions pertaining universally to Muslims. They do not reflect the more moderate conclusion that only some Muslims had any involvement in 9/11 whatsoever.

My Scottish grandmother captured the mood quite eloquently. When we spoke to her after the World Trade Center attacks, she weighed in with her opinion on the correct course for US foreign policy. In particular, she recommended that we should “jus’ drop a bomb on a’ tha’e Muslims.”⁵

The French/Algerian Conflict

The relationship between the French and the Algerians is a complicated one, as would be expected given their history of war, colonization and occupation. Anti-Algerian racism is still virulent in France, and North Africans have consistently been the victims of more hate crimes than any other group. The cognitive bias under investigation here is, of course, too simple to account for all the subtleties of a racism with such a complex

⁴ There a billion Muslims world-wide; very few of them are involved in any way with Al Qaeda, let alone with the specific acts of 9/11.

⁵ Translation: drop a bomb on all Muslims.

history. However, if we trace anti-Algerian prejudice to its early days, we find again the hand of generic reasoning.

The decade between 1910 and 1920 saw the first wave of significant Algerian immigration to France. During that period, the relationship between the French and the Algerian immigrants was generally quite friendly. An inquiry into the conditions of Algerian workers in 1914 found that relations between them and France's indigenous population were quite amicable (Beaugency 1914, reported in MacMaster 1995). In 1923, however, French-Algerian relations took a sharp turn for the worse. A wave of anti-Algerian violence began, in which North Africans were attacked at random. The attacks included a public lynching in the rue Fremicourt in Paris, and it was unsafe for North Africans to venture into the surrounding area. The media denigrated North Africans, and petitions were circulated that called for "the undesirables to be driven from the area" (MacMaster 1995: 150).

This overwhelming hostility towards North Africans was sparked by a single incident. On November 7 1923, Khemile Ousliman, an unemployed North African man, knifed a woman in the Rue Fondary. Ousliman, who was likely mentally ill, had been obsessed with the woman, and had repeatedly made sexual advances towards her. When she refused, he slit her throat, then turned in a frenzy on some passers by, killing another woman and wounding two others.

The French media quickly embellished the story. *L'Oeuvre* ran an article entitled "An Arab stabs four women", which described how a drunken "*Sidi* ... slit the shopkeeper open with an enormous cutlass ... [then] foaming at the mouth and brandishing his knife ... covered with blood, and his eyes bulging, he fell upon his new

prey [the second woman], letting out savage screams” (*L’Oeuvre* November 8 1923, reported in MacMaster 1995). Other papers discussed the “question of the ‘undesirable Algerians’,” and *L’Humanite* went as far as to describe how the local French were “terrorized by Arabs; they eat dogs, cats and rats. It’s not unlikely that they eat human flesh. If a crime, a rape, a theft or an attack is committed today or tomorrow, don’t hesitate: look for the Arab” (*L’Humanite*, November 9 1923, reported in MacMaster 1995).

So began a surge of anti-Arab discourse in the media, and anti-Arab violence and discrimination throughout France. Both the media’s sweeping generalizations and the random anti-Arab hate crimes were sparked by the actions of a single North African, who was most likely mentally ill. Seven years later, Paul Catrice, a Catholic priest and immigration expert, remarked that “If the *Sidi*, in general, inspires a certain repulsive fear, it is because of the memories of certain sensational crimes from which Parisians have drawn unconsidered generalizations” (MacMaster 1995: 158). These “unconsidered generalizations” are exactly the ones whose cognitive basis we are considering in this paper. There is no *a priori* reason to think that humans would be disposed to reason from a single sensational event to a category-wide generalization; certainly there is no logical demand for such thinking. We are unfortunately possessed of a particular bias that makes such reasoning not only possible, but prevalent.

Travel Logs and the Roots of Anglo-American Racism

Contemporary race-relations in America are again too complex to be assimilated to any single model of reasoning. If we look back far enough in history, however, I suggest that we see a role for generic reasoning in the formation of early anti-Black and anti-Native

American prejudices. To appreciate this, we have to look back almost 500 years, to the early days of English exploration.

In the mid-16th century, travel logs written by explorers were extremely popular. The population was rapidly becoming literate; in fact, it is estimated that about half of England was literate by 1533. Travel logs were embraced by the newly literate population, and played a significant role in the economic success of various print companies (Cole 1972: 60). Very few people could afford to travel abroad themselves, so the reports of a small number of explorers were the source of public knowledge of foreign lands and their inhabitants. The initial impressions of the English population vis-à-vis Blacks and Native Americans derived from these travel logs.

The initial interactions between English travelers and Africans were amicable, consisting primarily of the mutual trading of goods (Jordan 1968: 3). From the start, the English explorers were struck by the color of the Africans' skin, and the early travel logs devoted a significant amount of time to a discussion of just exactly how dark their skin was (Jordan 1968: 6-7; Cole 1972: 64). From early on, it was claimed that the dark color of their skin was innate and essential. In 1578, the voyager George Best claimed to have witnessed an Ethiopian settle in England and marry a white English woman. The couple had a child, who was "in all respects as blacke as the father was." Best concluded that "it seemeth this blacknes proceedeth rather of some natural infection of that man, which was so strong, that neither the nature of the Clime, neither the good complexion of the mother concurring, coulde any thing alter" (reported in Jordan 1968: 15). Such observations surely served to represent Blacks as an essentialized racial kind. There was something inherent and heritable in Black people that gave rise to their outward appearance.

It did not take long for the explorers to start to produce sensational reports.

Winthrop Jordan, in his discussion of travel logs on Africa, writes

To judge from the comments of voyagers, Englishmen had an unquenchable thirst for the details of savage life ... It is scarcely surprising that civilized Englishmen should have taken an interest in reports about cosmetic mutilation, polygamy, infanticide, ritual murder and the like – of course *English* men did not really *do* any of these things themselves ... It would be a mistake to slight the importance of the Negro's savagery, since it fascinated Englishmen from the very first. English observers in West Africa were sometimes so profoundly impressed by the Negro's deviant behavior that they resorted to a powerful metaphor with which to express their own sense of difference from him. They knew perfectly well that Negroes were men, yet they frequently described the Africans as "brutish" or "bestial" or "beastly." The hideous tortures, the cannibalism, the rapacious warfare, the revolting diet (and so forth page after page) seemed somehow to place the Negro among the beasts (Jordan 1968: 25-28).

These travel logs, which did so much to shape England's early image of Africa, contained endless gory accounts of shocking behavior on behalf of the Africans. At the same time, they also carried a powerful implication that Blacks were to be located among the non-human animals. While this carries some obvious implications – that they are sub-human and so need not be treated with normal human dignity – our discussion suggests that the comparison may even have other, more subtle psychological effects. Since basic-level animal kinds are essentialized kinds *par excellence*, once 'the Negro' is represented as a comparable kind, the mechanism of primitive generalization will treat 'the Negro' as an appropriate locus for a striking property generalization, in much the same way that the shark or the mosquito is.

This fatal combination of essentialization and lurid descriptions of negative and striking behavior on behalf of Africans exploits the original, simplifying sin of cognition. Thus the sensational logs led quickly to the familiar generalizations and the prejudices

they engender. And these travel logs, it must be remembered, were the primary source of information for the English in these early days of intercontinental exploration.

It should be noted that the travelers themselves often reported only specific incidents of cannibalism, or other specific instances of horrific violence. That is, it would be overly simplistic to place the blame for the formation of early negative stereotypes squarely on the explorers. Many of them were quite responsible in their reporting, and did not indulge themselves in broad generalizations. The rather depressing upshot of the nature of our default generalizations is that they did not have to. The reporting of specific instances would suffice to form very general beliefs in the mind of the reader.

It also must be noted that these travel logs recorded a large amount of neutral or fairly positive information. One must look through many pages of descriptions of foods, spices, headdresses, and even some quite charming customs before one finds descriptions of appalling events. As we would predict, however, it was the degree of horror the descriptions inspired, not their frequency or even generality, that determined the extent to which they shaped people's thinking.

The Essentialization of Racial, Ethnic and Religious Kinds

As noted above, Africans were compared to beasts from the outset, which may well have contributed to their being viewed as an essentialized kind, on a par with non-human basic-level kinds. This perception was fueled in part by the discovery of the 'orang-outang' (in fact it was the chimpanzee) alongside Africans. English explorers encountered orang-outangs and Africans at the same time, and the coincidence of discovering this most human-like animal in the same time and place as dark-skinned Africans aroused speculations. Some suggested that Blacks were the offspring of orang-

outangs, and others that orang-outangs were the offspring of Blacks and unknown beasts. These theories were not too widely accepted; however, many people did become convinced that male orang-outangs would copulate with African women. This rumor, a version of which is endorsed by none other John Locke, persisted into the 18thC (Jordan 1968: 28-32). To the extent that Blacks were associated with orang-outangs in people's minds, this could only have exacerbated the perception of Blacks as possessed of an essential nature that was fundamentally different from that of Europeans.

There were numerous 'scientific' attempts to investigate these essential differences. The study of phrenology, for example, consisted in large part of attempting to find differences in the respective skulls of black and white subjects. The Swedish naturalist, Carl Linnaeus, developed an intricate and sophisticated taxonomy of species, but unfortunately located 'the Negro' between apes and Europeans in his classification. In the centuries that followed Linnaeus' work, much effort was devoted to attempts to scientifically discover the essential differences between the races. Nancy Stepan writes of this:

Closely related to abstraction was the typological orientation scientists brought to their studies of race. To the typologist, every individual human being belonged in some way or another to an undying essence or type. However disguised or hidden the individual's membership in the type might be, the scientists expected to be able to see behind the individual to the type to which he belonged. The result was to give a 'mental abstraction an independent reality' to make real or 'reify' the idea of racial type when in fact the type was a social construct which scientists then treated as though it were in fact 'in nature' (Stepan 1982: xviii).

The orientation that Stepan attributes to the scientists in question reflects the essentialization of racial kinds. Primitive biological science may have helped to crystallize such beliefs, but they may well have existed at the intuitive level of folk

biology, just as similar beliefs concerning animal kinds can be found in untutored pre-schoolers.

Nowadays, justifying racist generalizations on biological grounds is thankfully on the back foot in academic circles, even if it persists among laypeople. Unfortunately, the same cannot be said for generalizations that range over religious groups such as Muslims.

In an influential article entitled “The Roots of Muslim Rage,” emeritus Princeton professor Bernard Lewis argues that Islam is an inherently violent religion. He points out that Mohammed was “not only a prophet and a teacher, like the founders of other religions; he was also the head of a polity and of a community, a ruler and a soldier” (Lewis 1990). He argues that Islam divides the world into two houses: the House of Islam, and the House of War. We are now facing nothing less than a ‘clash of civilizations,’ since “Islamic fundamentalism has given an aim and a form to the otherwise aimless and formless resentment and anger of *the Muslim masses*” (Lewis 1990, my emphasis – note the high level of generality implied by his language). Lewis goes on to describe Islam and its followers in generic, essentialized terms:

There is something in the religious culture of Islam which inspired, in even the humblest peasant or peddler, a dignity and a courtesy toward others never exceeded and rarely equaled in other civilizations. And yet, in moments of upheaval and disruption, when the deeper passions are stirred, this dignity and courtesy toward others can give way to an explosive mixture of rage and hatred with impels even the government of an ancient and civilized country – even the spokesman of a great spiritual and ethical religion – to espouse kidnapping and assassination, and try to find, in the life of their Prophet, approval and indeed precedent for such actions (Lewis 1990).

The distinguished professor is saying that in virtue of following Islam, the arbitrary Muslim– be they peasant or peddler – is disposed in times of calm to be most courteous towards other, but when faced with upheaval, this person’s finer dispositions give way to

an “explosive mixture of rage and hatred.” Lewis implies that there is just *something about Islam* that affects people in this way. In virtue of being Muslim, people possess the dispositions he describes. This is a way of essentializing the followers of Islam, of attributing a shared essence to them that causally grounds their putative violent dispositions. In effect, Lewis is offering a rationalization of anti-Muslim generalizations.

I would suggest that attempts to essentialize kinds in such a way as to support general attributions of violent tendencies to their members are *post hoc* attempts to consciously justify conclusions delivered by primitive cognitive biases.

Given that we have a disposition to unthinkingly essentialize racial and related social kinds, we are vulnerable to making striking property generalizations over them. These generalizations, blindly made in the mind’s intuitive depths, rise to consciousness as deep-rooted convictions. It is only then, at this late point in the process, that we look around for intellectual justification for these judgments. Being a creative species, we are able to come up with fanciful tales that range from the African’s sexual involvement with orang-outangs, to the Muslim’s inherent inclination towards explosive rage and hatred.

It is just storytelling.

Looking for Hope

It is normally quite satisfying to think that one’s theory may be correct, but I find no satisfaction in the thought that I might be near the truth in this matter. The discussion indicates that racist beliefs have their roots deep in our cognitive structure. They are formed by our default mechanism of generalization; our innate capacity to move from the particular to the general is skewed in such a way as to open the door to prejudice. We have a deep-rooted way of thinking that lets us learn quickly to avoid any sharks that

venture near our beaches, to keep our children out of the way of pit-bulls, and to meticulously check for deer ticks after a day in the woods – and what evolutionary benefits must inhere in this way of generalizing! But along with the selective advantage conferred by this way of thinking came the clear path to prejudice. In this basic cognitive bias of ours lies cognition's original sin.

In the early days of combating racism, it was pointed out that so many of the racist stereotypes people held were certainly not true of *all* the members of the minority, and in fact were not true of even *most*. These observations did not lead to the widespread rejection of those stereotypes. That this would be so is not surprising in light of our discussion. Just as the generalization that sharks attack bathers can be retained in full knowledge that most of them do not, racist generalizations can be impervious to statistics. The suspicion of the general underlying negative disposition lingers even in the face of statistically low levels of its manifestation.

How, then, might we try to combat prejudice in light of these observations? One hopeful observation is the tendency of people who are very familiar with a class to reject striking property generalizations over that class. People with considerable experience with Catholic priests tend to reject the generic claim that priests molest children. This claim is far more likely to be made by someone with only the most limited contact with Catholic priests. As a matter of sad fact, the incidence of molestation by fathers is significantly higher than by priests, and we are none of us inclined to accept the generalization that fathers molest children.

What is going on here? It seems that familiarity with members of a kind helps inhibit the tendency to make striking property generalizations over the kind. Groups of

which we are members, or which we have detailed personal knowledge, are less likely to be targets of this particular sort of generalization. Of course we may well be aware that *some* members of the familiar group may have committed some or other horrific deed, but we do not generalize this information in the way that we do when dealing with members of less familiar groups. The difference, I propose, lies in the differential tendency to view the deplorable actions as grounded in the nature or essence of the group in question. If a member of a group with which we identify commits an appalling act, we do not reach beyond that individual in attributing the disposition to so act. If such an individual belongs to a group with which we do not identify, however, we may well view the inclination to appalling action as part and parcel of the essence of that group.

If this hypothesis is correct, we have identified a specific sort of “attribution error.” Long ago, psychologists identified the so-called Fundamental Attribution Error, which consists in a certain asymmetry between attributions to oneself and to others. People tend to attribute their own less-than-stellar behavior in a given situation to aspects of that situation, and yet attribute identical behavior in others to persistent personality traits. For example, we may attribute another’s involvement in a car accident to her being a bad driver, but would attribute our own involvement to external factors such as poor weather conditions (Jones and Nisbett 1972). There has been some discussion of another attribution error, known as the Ultimate Attribution Error (UAE), which is concerned with attributions across groups (Pettigrew 1979). To be guilty of the UAE, one must be inclined to attribute negative actions of members of groups to which we belong (in-groups) to situational factors, but attribute the negative actions of members of groups to which we do not belong (out-groups) to persistent traits. There is some empirical

evidence that we are susceptible to the UAE, but it is less than overwhelming (see Hewstone 1990 for a review).

The attribution error that I have in mind should not be confused with the Ultimate Attribution Error. The UAE would have us explain away negative actions of in-group members by reference to situational factors in the way we explain away our own bad behavior, while attributing negative actions of out-group members to more persistent factors. The UAE thus offers the same contrast as the Fundamental Attribution Error. The attribution error I have in mind here – perhaps we should call it the Supreme Attribution Error, so as not to be outdone – is concerned with a different contrast. We may very well explain negative behavior of in-group members by reference to their persistent personality traits, but *we will not explain it by reference to the group essence*. In contrast, when faced with highly negative behavior on the part of members of unfamiliar essentialized groups, we may take the disposition to such behavior to belong not only to the personality of perpetrator, but to the very nature or essence of the group.

To illustrate the difference between the UAE and the attribution error I am discussing here, consider the reactions on behalf of Americans to the Oklahoma City bombings, and the bombings of 9/11. To commit the UAE here, we would have to attribute Timothy McVeigh's actions to situational factors, while refraining from offering such excuses for the 9/11 bombers. I'm sure we have very little inclination to think that McVeigh was simply a victim of circumstances. By contrast, to commit the attribution error that I am describing, one can believe, as we probably do, that McVeigh's actions were wholly the consequence of his rather demented personality. The attribution error would lie in understanding McVeigh's dispositions to be grounded in his individual

personality, while taking the 9/11 bombers to be manifesting an ‘explosive mixture of rage and hatred’ inherent in Muslims.⁶

How do we then cope with sensational negative information concerning members of such familiar kinds? The case of dogs is telling. When a Rottweiler mauls a child, we do not impugn dogs in general. We instead seek a more restrictive generalization that cites a better predictor of the tendency to maul children. While a Labrador owner might rest content with the generalization that Rottweilers maul children, the Rottweiler breeder will further restrict the generalization to, say, *poorly trained* Rottweilers. The intellectual tendency to look for better predictors can be generalized to create an appropriate skepticism toward our reflex generic thoughts and remarks.

As noted earlier, familiarity with counterexamples is not in itself enough to defeat the force of striking property generalizations. A key factor seems to be a significant degree of *identification* or *solidarity* with members of the kind. If true, this would expose the fatal, reinforcing “logic” of segregation. The more separated we are, the more see each other in the terms that invite invidious striking property generalizations.

This observation fits nicely with the so-called contact hypothesis, originally advanced by G. W. Allport in 1954. He predicted that contact between members of

⁶ This attribution bias should not be confused with the out-group homogeneity bias (Quattrone and Jones 1980; Wood, Jones and Quattrone 1981; Park and Rothbart 1982). The out-group homogeneity bias consists in people’s perceiving out-groups to be more homogenous than in-groups. The magnitude of this effect – while robust and statistically significant – is not large; we are inclined to rate out-groups as only slightly more homogenous than in-groups (Mullen and Hu 1989). It does not entail our perceiving an out-group to be such that all its typical members share a given disposition. The out-group homogeneity bias is also concerned with a wide range of traits, including (often primarily) positive and neutral ones. It is not at all limited to strikingly negative ones; in fact, if a study of perceived out-group homogeneity that focuses on strikingly negative characteristics exists, I have not found it.

There is also evidence that members of minority groups may exhibit the reverse bias, rating themselves as *more* homogenous than the majority out-group (Guinote 2001; Simon and Brown 1987; Simon and Pettigrew 1990). I would not expect that minority members would be inclined to make these striking property generalizations over their own groups, however. The out-group homogeneity bias is a well-documented and important bias, but it does not explain the attribution bias I am concerned with here.

different groups would reduce prejudices directed towards the other group, but *only* if a) the groups have equal status, b) they participate in cooperative activity towards common goals, c) the acquaintance is personalized, and d) the contact is sanctioned by authorities and/or social norms. Whether all these conditions must be met, and whether there are additional necessary conditions has been a matter of some controversy over the ensuing decades, but the core idea has received considerable empirical support (Pettigrew and Tropp 2000). Prejudice is reliably reduced when members of different groups interact in cooperative and personal ways. The conditions that the contact hypothesis sets forth may well be conditions that trigger our mechanism to seek for a better predictor of striking properties.

Other Psychological Explanations of Prejudice

It should be clear that the thesis set out here need not conflict with other theories of the psychological roots of racism. I rather take my thesis to be quite compatible with them, allowing that each theory may delineate a different aspect of this complex phenomenon. For example, the sophisticated analysis of group identification provided by Social Identity Theory (e.g. Tajfel 1982, Tajfel and Turner 1986) is extremely important to our understanding of group dynamics, and is not at all challenged by my identification of this particular bias of ours. Similarly, Realistic Conflict Theory's insights (e.g. Sherif 1966) into the role of competition over limited resources fill in an important part of the picture that my account does not touch. It would be quite surprising if prejudice turned out to have a single and uniform psychological basis, rather than being the result of many disparate factors. To seek *the* psychological explanation for prejudice is likely a mistaken quest.

However, it is important to distinguish my thesis from the thesis that racist generalizations are the result of illusory correlations (Chapman 1967, Hamilton and Gifford 1976, Mullen and Johnson 1990). Illusory correlations – a well supported phenomenon – occur when people overestimate the degree to which two rare events co-occur. If event A is rare relative to event B, and event C is rare relative to event D, then people may guess that A and C co-occur far more often than B and D, even though the actual rates of co-occurrence are the same. Hamilton and Gifford (1976) were the first researchers to propose that racist generalizations like the ones discussed here may be explained by illusory correlations – it is more rare to encounter a member of a minority than a member of the majority, and strikingly awful events are, thankfully, more rare than neutral ones. Perhaps we are susceptible to illusory correlations in such cases, and so overestimate how often minority members commit horrific acts.

Illusory correlations are a real phenomenon, and I do not doubt that they play an unhelpful role in the formation of prejudices. I believe, however, that they are insufficient to explain the nature of the generalizations discussed here. It's important to note that, while two rare events may be judged to co-occur more often than two more common events, the magnitude of the difference is not huge. It is certainly not the case that the two rare events are taken to co-occur *all* the time, or even *most* of the time. Illusory correlations would predict that the Ingall's neighbor Mrs. Smith might overestimate the number of Indians that have been involved in massacres, but would not predict the sweeping universality of her claim "the only good Indian is a dead Indian". The scope of these generalizations goes far beyond what can be reasonably predicted by illusory correlations.

Further, while the experimental support for illusory correlations is not at all limited to striking or negative events – illusory correlations can be observed between two perfectly neutral or positive events (see, e.g., Hamilton and Gifford 1976 experiment 2) – theorists tend to only employ them in explaining negative stereotypes, often ones involving strikingly negative behaviors of the sort here discussed. One tends to find that such a move is justified by the most cursory of comments; Hamilton and Gifford note, for example, that “since for most varieties of behavior the norm is positive in value, undesirable (non-normative) behavior is statistically less frequent than desirable behavior and [so] can also be considered distinctive” (1976, p. 394) If we take illusory correlations seriously as an explanation of stereotype formation, though, we must predict that stereotypes will form whenever a minority member engages in a rare activity. It should only be the relative rarity of the group and the activity that account for whether the stereotype is formed, not the nature of the activity itself. I suppose that bouncing around on a pogo stick is a very rare activity. Speaking from personal experience, I’ve certainly heard tell of far more instances of terrorism than I have of people bouncing around on pogo sticks. Assuming my statistical estimate is relatively accurate – and if it’s not, we could substitute another comparable example – we would have to predict that we would be at least as inclined to conclude that *Muslims bounce around on pogo sticks* upon hearing of a few Muslims engaging in this activity as we were to conclude that *Muslims are terrorists* in the wake of 9/11.

Again, this is not to deny that illusory correlations are a real phenomenon, or that they play some role in our prejudicial attitudes, but rather to point out that they alone cannot explain the generalizations with which I am concerned here.

Conclusion: Combating these Generalizations

Recently, Anthony Appiah, in a humorous article in *Slate*, when asked what he would propose as the non-negotiable core of the university curriculum, suggested two things: junior year abroad, and Bayes' Theorem. Indeed, it is personal experience with members of other racial and ethnic groups, along with a quasi-statistical tendency to look for better predictors that may be our best bulwark against the original sin of cognition.

Raising awareness of the original sin may also be helpful. (When someone offers a striking property generalization, ask them whether they can find a better predictor.) The habit of mind would at least make us wary of attempts such as Bernard Lewis' to rationalize striking property generalizations. We should wonder that whatever plausibility they have may derive from this bias of ours, rather than from sources governed by reason. Of course it would have been more helpful to have been aware of this bias five hundred years ago, when so many prejudices were in their infancy. Half a millennium of politics and power have complicated these prejudices; it remains an open question whether the clarifying power of cognition can correct them.

References

Baldwin, D. A., E. Markman, and R. Melartin. (1993). "Infants' Ability to Draw Inferences about Nonobvious Object Properties: Evidence from Exploratory Play". *Child Development*, vol. 64, pp. 711-728.

Chapman, L. J. (1967). "Illusory Correlation in Observational Report". *Journal of Verbal Learning and Verbal Behavior*, vol. 6, pp. 151-155.

Cole, R. (1972). "Sixteenth Century Travel Books as a Source of European Attitudes toward Non-White and Non-Western Culture". *Proceedings of the American Philosophical Society*, vol. 116, no. 1, pp. 59-67.

Fara, M. (2005). "Dispositions and habituals". *Nous*, 39, pp. 43-82.

Gelman, S. A. (2003). *The Essential Child*. New York: Oxford University Press.

Gelman, S. A., & Wellman, H. M. (1991). „Insides and essences: Early understandings of the nonobvious”. *Cognition*, 38, pp. 213–244.

Graham, S. A., C. S. Kilbreath, and A. N. Welder. (2001). “Words and Shape Similarity Guide 13-month-olds’ Inferences about Nonobvious Object Properties”. In J. D. Moore and K. Stenning (eds.), *Proceedings of the Twenty Third Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum, pp. 352-357.

Guinote, A. (2001). “The Perception of Group Variability in a Non-Minority and a Minority Context: When Adaptation Leads to Out-group Differentiation”. *British Journal of Social Psychology*, vol. 40, pp. 117-132.

Hamilton, D. L. and Gifford, R. K. (1976). “Illusory Correlation in Interpersonal Perception: A Cognitive Basis of Stereotypic Judgment”. *Journal of Experimental Social Psychology*, vol. 12, pp. 392-407.

Hewstone, M. (1990). “The ‘Ultimate Attribution Error’? A Review of the Literature on Intergroup Causal Attribution”. *European Journal of Social Psychology*, vol. 20, pp 311-335.

Hirschfeld, L. A. (1996). *Race in the Making*. Cambridge, MA: MIT Press.

Hollander, M. A., S. A. Gelman, and J. Star. (2002). “Children’s Interpretation of Generic Noun Phrases”. *Developmental Psychology*, vol. 36, no. 6, pp. 883-894.

Jones, E. E. and Nisbett, R. E. (1972). “The actor and the observer: Divergent perceptions of the causes of the behavior.” In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins and B. Weiner (eds.), *Attribution: Perceiving the causes of behavior*. Morristown, NJ: General Learning Press, pp. 79-94.

Jones, E. E., Wood, G. C. and Quattrone, G. A. (1981). “Perceived Variability of Personal Characteristics in In-groups and Out-groups: The Role of Knowledge and Evaluation”. *Personality and Social Psychology Bulletin*, vol. 7, pp. 523-528.

Jordan, W. (1968). *White over Black: American Attitudes toward the Negro, 1550–1812*. Chapel Hill: University of North Carolina Press.

Leslie, S. J. (2007). “Generics and the Structure of the Mind”. *Philosophical Perspectives*, pp. 375-405.

Leslie, S.J. (2008). “Generics: Cognition and Acquisition”. *The Philosophical Review*, vol. 117, no. 1, pp. 1-49.

Lewis, B. (1990). "The Roots of Muslim Rage". *The Atlantic Monthly*; September 1990, vol. 266.

Lewis, D. (1997). "Finkish dispositions". *Philosophical Quarterly*, 47, pp. 143-58.

MacMaster, N. (1995). "The Rue Fondary Murders of 1923 and the Origins of Anti-Arab Racism," In Jan Windebank and Renate Gunther (eds.), *Violence and Conflict in the Politics and Society of Modern France*. Lampeter, UK: Edwin Mellen Press, pp 149-160.

Medin, D. L., & Ortony, A. (1989). "Psychological essentialism". In S. Vosniadou & A. Ortony (eds.) *Similarity and Analogical Reasoning*. New York: Cambridge University Press, pp. 179-196.

Mullen, B. and Hu, L. (1989). "Perceptions of Ingroup and Outgroup Variability: A Meta-Analytic Intergration". *Basic and Applied Social Psychology*, vol. 10, no. 3, pp. 233-252.

Mullen, B. and Johnson, C. (1990). "Distinctiveness Based Illusory Correlations and Stereotyping: A Meta-Analytic Integration". *British Journal of Social Psychology*, vol. 29, pp. 11-28.

Park, B. and Rothbart, M. (1982). "Perception of Out-group Homogeneity and Levels of Social Categorization: Memory for the Subordinate Attributes of Ingroup and Outgroup Members". *Journal of Personality and Social Psychology*, vol. 42, pp. 1051-1068.

Pettigrew, T. (1979). "The Ultimate Attribution Error: Extending Allport's Cognitive Analysis of Prejudice". *Personality and Social Psychology Bulletin*, vol. 5, no. 4, pp. 461-476.

Pettigrew, T. and Tropp, L. (2000). "Does Intergroup Contact Reduce Prejudice? Recent Meta-Analytic Findings". In S. Oskamp (ed.) *Reducing Prejudice and Discrimination*. Mahwah, N.J.: Lawrence Erlbaum Associates, pp. 93-115.

Quattrone, G. A. and Jones, E. E. (1980) "The Perception of Variability Within Ingroups and Outgroups: Implications for the Law of Small Numbers". *Journal of Personality and Social Psychology*, vol. 38, pp. 141-152.

Rosch, E. H. (1978). "Principles of categorization". In E. Rosch & B. Lloyd (eds.), *Cognition and Categorization*. Hillsdale, N.J.: Erlbaum Associates, pp. 27-48.

Rothbart, M., & Taylor, M. (1992). "Category labels and social reality: Do we view social categories as natural kinds?" In G. Semin and K. Fiedler (eds.), *Language, Interaction and Social Cognition*. London: Sage, pp. 11-36.

Simon, B. and Brown, R. (1987). "Perceived Intra-group Homogeneity in Minority-Majority Contexts" *Journal of Personality and Social Psychology*, vol. 53, pp. 703-711.

Simon, B and Pettigrew, T. (1990). "Social Identity and Perceived Group Homogeneity: Evidence for the Ingroup Homogeneity Effect". *European Journal of Social Psychology*, vol. 20, pp. 269-286.

Sherif, M. (1966) *In common predicament: Social psychology of intergroup conflict and cooperation*, Boston: Houghton-Mifflin

Stepan, Nancy. (1982). *The idea of race in science*. London: Macmillan.

Tajfel, H. (1982). *Social identity and intergroup behavior*. Cambridge, England: Cambridge University Press.

Tajfel, H. and Turner, J. C. (1986). "The Social Identity Theory of Intergroup Behavior". In S. Worchel and W. Austin (eds.) *Psychology of Intergroup Relations*. Chicago: Nelson Hall.