

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

Penultimate Draft: Please refer to final version for citation purposes

Introduction

Long before we learn to talk, our expectations concerning novel members of a category are shaped by our experience with already encountered members. We expect, for example, that objects that share obvious perceptible qualities will also share dispositional properties. If a given item rattles when shaken, nine-month olds expect that other items which share the same perceptible profile will rattle when shaken.¹ 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’.² From the very beginning, we are inclined 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. In earlier papers, I argue that generics – sentences such as ‘ravens are black’ and ‘tigers are striped’ – express the generalizations that are delivered by

* I would like to thank the following people for their invaluable help, comments, and conversation: K. Anthony Appiah, Andrei Cimpian, Kai-Yuan Cheng, Susan Gelman, Tamar Szabo Gendler, Sam Glucksberg, Susan Fiske, Elizabeth Harman, Sally Haslanger, Richard Holton, Mark Johnston, Thomas Kelly, Sangeet Khemlani, Rae Langton, Eden Lin, Ron Mallon, Sally McConnell-Ginet, Sarah McGrath, Alexander Nehamas, Joyce Carol Oates, Rachel Parsons, Debbie Prentice, Mark Richard, Gideon Rosen, Sandy Waxman, Helen Yetter, the members of my fall and spring 2009 *Philosophical Issues in Language and Cognition* class, the audience at the New York Society for Women in Philosophy, and anonymous *Journal of Philosophy* reviewers.

¹ D. A. Baldwin, E. Markman, and R. Melartin, “Infants’ Ability to Draw Inferences about Nonobvious Object Properties: Evidence from Exploratory Play”, *Child Development*, 64 (1993): 711-728.

² S. A. Graham, C. S. Kilbreath, and A. N. Welder, “Words and Shape Similarity Guide 13-month-olds’ Inferences about Nonobvious Object Properties”, *Proceedings of the Twenty Third Annual Conference of the Cognitive Science Society*, (2001): 352-357.

this basic mechanism of generalization.³ If this is so, then generics provide us with a window onto the workings of this mechanism. In this paper, I am concerned with a particular aspect of this mechanism, namely the route by which we reach general conclusions regarding dangerous or harmful features. What follows in no way purports to be an exhaustive analysis of all the factors that have formed and sustained prejudiced attitudes. Rather, my aim is to identify and discuss *one* particular cognitive bias that has given birth to many a prejudice.

While I adopt a cognitive perspective here, this is not to imply that economic, political and cultural perspectives are not of equal (or even greater) value and importance. These various perspectives are not in competition with each other; rather they complement each other by providing different levels of explanation. Moreover, even within the domain of cognitive explanations of prejudice, I do not purport to offer anything close to a full psychological account of prejudiced attitudes. The focus is on a particular subset of negative stereotypes: ones that involve generalizing extreme and horrific behavior from a few individuals to a group, for example *Muslims are terrorists* or *Blacks are rapists*.

Two things, though obvious, are worth noting at the beginning. First, cognitive bias explanations do not excuse racial or cultural prejudice, any more than noting that we are hardwired to seek out and accumulate resources serves as an excuse for extreme covetousness or for theft. Secondly, offering a psychological explanation for prejudice does *not* entail that prejudice is inevitable. Quite the contrary – the closing sections of this paper will discuss some ways in which we might combat prejudice, with a particular focus on how we might *prevent* the formation of these attitudes in the course of childhood development. These suggestions are based on recent psychological research, and so – far from implying

³ S. J. Leslie, “Generics: Cognition and Acquisition”, *Philosophical Review*, 117, 1, (2008): 1-49; S. J. Leslie, “Generics and the Structure of the Mind,” *Philosophical Perspectives*, (2007): 375-405.

that prejudice is an inevitable feature of human psychology –the cognitive perspective on prejudice may point to some novel means of combating it.⁴

Striking Property Generalizations

In previous work, I argue that a variety of philosophical, linguistic and psychological considerations suggest that generic sentences may be language’s way of letting us give voice to cognitively primitive generalizations.⁵ This hypothesis has subsequently received further support by new psychological data.⁶ There are now a variety of convergent reasons for supposing that the generalizations we articulate using generics reflect deep-seated aspects of our psychology.

In theorizing about generic generalizations it is helpful to identify various sub-classes of these generalizations, one of which I term ‘striking property’ generalizations. This class includes claims such as:

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

These claims 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,’

⁴ For an exemplary defense of the utility of psychological accounts of prejudice against various criticisms and misunderstandings, see E. Machery, L. Faucher, L., and D. Kelly, “On the Alleged Inadequacy of Psychological Explanations of Racism”, forthcoming in *The Monist*.

⁵ Leslie, 2008, *op. cit.*; Leslie, 2007, *op. cit.*

⁶ See, e.g., S. J. Leslie, S. Khemlani & S. Glucksberg, “All Ducks Lay Eggs: The Generic Overgeneralization Effect”, *under review*; S. J. Leslie & S. A. Gelman, “Quantified Statements are Recalled as Generics: Evidence from Preschool Children and Adults”, *under review*; A. Brandone, A. Cimpian, S. J. Leslie, and S. A. Gelman, “Do Lions have Manes? Children Interpret Generics as Making Claims about Kinds rather than Quantified Sets”, *under review*.

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 may appear that these generics require for their truth only that *some* of the kind possess the property in question.⁷ This is not true for generics in general; for example, *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 strong interest in avoiding.⁸ If even just a few members of a kind possess 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 give voice to psychologically primitive generalizations, this observation implies that our basic way of dealing with dangerous or harmful information involves the rapid generalization of this information to the appropriate kind or category. We do not wait around to see 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

⁷ 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, 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.

⁸ For some empirical confirmation that generics are more likely to be accepted at low prevalence levels if the property in question is dangerous, see A. Cimpian, A. Brandone, and S. Gelman, “Generic statements require little evidence for acceptance but have powerful implications”, forthcoming in *Cognitive Science*.

potentially huge. Our ancestors were far better off jumping to conclusions, as it were, rather than taking the time 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 requires one to worry with considerable regularity, whereas a single murder suffices to make one a murderer.

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.⁹

The ‘introduction’ conditions, as it were, of striking property generalizations – how the world must be for us to form or accept these generalizations – are very undemanding when it comes to how prevalent the property has to be in the relevant population. What, though, of the ‘elimination’ conditions of these generalizations – how does acceptance or rejection of these generalizations impact the inferences we are willing to draw?¹⁰ We are content to accept “ticks carry Lyme disease” despite knowing that very few ticks actually carry the relevant bacterium. One might suppose that an ideally rational agent would be very

⁹ 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.

¹⁰ I must emphasize that I mean “elimination conditions” here to be read as wholly psychological, and not at all normative (hence the scare quotes). That is, I mean to highlight the inference that we *actually* draw from these generalizations, not the inferences that we *ought* to draw from them. For more discussion, see S. J. Leslie, *Generics*, (Oxford: Oxford University Press, forthcoming).

hesitant to suppose that an arbitrary tick carries Lyme disease, in light of these statistical facts.

However, recent psychological results suggest that acceptance of a generic strongly influences our judgments concerning whether an arbitrary member of a kind has a property *over and above our beliefs about the prevalence of the property*. This is especially so for striking property generalizations. Sangeet Khemlani, Sam Glucksberg and I found that people were as likely to agree that *Jumpy the tick carries Lyme disease* as they were to agree that *Joe the Canadian is right-handed* – despite the very large discrepancy between the subjects own (roughly correct) judgments of the prevalence of the respective properties in the respective populations. Sixty-five percent of our participants who accepted the striking property generalizations judged – with varying degrees of confidence – that an *arbitrary* member of the kind would have the striking property.¹¹ Andrei Cimpian, Amanda Brandone, and Susan Gelman found comparable results using a very different experimental design. While their participants frequently accepted novel striking property generics at low prevalence levels, if they were presented with a novel striking property generic and asked to

¹¹ We showed participants the following information:

Suppose you are told: Jumpy is a tick.

What do you think of the following statement: Jumpy carries Lyme disease.

Participants responded by marking a 7-point scale that ranged from “strongly agree” to “I couldn’t possibly tell” to “strongly disagree”. Participants were never shown a generic in this portion of the experiment, so they had to rely solely on background beliefs. A variety of examples were used, with varying degrees of prevalence and predictiveness (cue validity). Acceptance of the generic was found to predict responses above and beyond estimates of prevalence and predictiveness. As an intuitive illustration of this, average prevalence estimates for striking property generalizations were 33% (already inflated relative to actual fact), and average prevalence estimates for items like “Canadians are right-handed” were 60%, yet the mean responses were the same for both types of items despite the gross differences in average estimated prevalence (and overall similar ratings of cue validity). For more details, see S. Khemlani, S. J. Leslie & S. Glucksberg, “Generics, Prevalence, and Default”, *Proceedings of the 31st Annual Cognitive Science Society* (2009).

estimate how prevalent the property might be among the kind, they gave extremely high estimates – in many cases, 100%.¹²

These findings suggest that these generalizations play a powerful role in guiding our inferences concerning property possession, despite their relatively weak acceptance conditions. These primitive generalizations are not psychologically inert – rather they play a powerful role in guiding our judgments about members of a kind. In an early paper on generics, Robert Abelson and David Kanouse noted that some generics require very little evidence for acceptance, and yet “once accepted psychologically they appear to be commonly taken in a rather strong sense, as though the quantifier *always* had implicitly crept into their interpretation”.¹³ Our most primitive method of generalization seems to encourage us in reasoning from ‘some’ to “many” or “most”, or even to “all”, at least when striking properties are in play.

Fear, Prejudice, and Generalization

The cognitive disposition to generalize strikingly negative information very widely may serve a useful purpose in the non-social realm. When we turn to generalizations about groups of people, however, it can lead to disastrous consequences. I should reiterate that what follows is in no way intended as an exhaustive account of the cognitive factors underlying racism, and it does not touch on the social and cultural factors that have bred and sustained

¹² Cimpian et al., *op. cit.* Participants who were told that, e.g., *10%, 30%, or 50% of lorches have feathers that can cause massive bleeding* frequently accepted the generic “lorches have dangerous feathers”. However, participants who were instead told “lorches have dangerous feathers”, and then asked to estimate the prevalence, gave much higher estimates – often as high as 100%. Interestingly, this asymmetry was only found if the property was striking and/or ‘characteristic’ of the kind. For more accidental properties such as *having muddy feathers*, no such asymmetry was found. See Leslie, 2008, *op. cit.* for a discussion of the different types of generics – striking, characteristic, and more the ‘accidental’ majority generics.

¹³ Interestingly, their clearest examples are of striking property generics, though they do not identify them as such. (They do, however, note that it is the predicate that determines how likely the generic is to be accepted on the basis of weak statistical evidence.) R. P. Abelson & D. E. Kanouse, “Subjective acceptance of verbal generalizations.” In S. Feldman (Ed.), *Cognitive consistency: Motivational antecedents and behavioral consequents* (New York: Academic Press, 1966: 171-197), here p. 172; see also D. Schneider, *The Psychology of Stereotyping* (New York: Guilford Press, 2004).

racism. Rather, I seek only to identify a particular and pervasive aspect of our thinking which itself leads us down the dark road to prejudice. Our most primitive method of generalization has the potential to enshrine a pervasive form of bigoted thinking.¹⁴

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, ethnic, or religious minority (or “out-group”) may lead us to a general belief concerning their entire group. I discuss below why it may be that these particular social groups, as opposed to others, tend to be the targets of such generalizations, and I then identify the fundamental error behind these generalizations in the case of social groups. (That is, I explain how it may be true that *mosquitoes carry the West Nile virus*, and yet decidedly false that, say, *Muslims are terrorists*.) Once we identify the enabling error behind such generalizations over social groups, the opportunity arises to consider novel ways of combating this way of thinking.

As a result of a profound and pervasive cognitive bias built into our most primitive method of generalization, a few appalling acts on behalf of some members of a given group can lead us to draw conclusions about the group in general. As the available experimental evidence suggests, acceptance of a striking property generalization can lead one to draw the

¹⁴ 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 is extremely important to our understanding of group dynamics, and is not at all challenged by my identification of this particular cognitive bias of ours (e.g. H. Tajfel, *Social identity and intergroup behavior* (Cambridge, England: Cambridge University Press, 1982); H. Tajfel & J. C. Turner, “The Social Identity Theory of Intergroup Behavior”, in S. Worchel and W. Austin (eds.) *Psychology of Intergroup Relations* (Chicago: Nelson Hall, 1986)). Similarly, Realistic Conflict Theory’s insights into the role of competition over limited resources fill in an important part of the picture that my account does not touch (e.g. M. Sherif, *In common predicament: Social psychology of intergroup conflict and cooperation* (Boston: Houghton-Mifflin, 1966)). 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.

corresponding conclusion about the arbitrary member of the group – conclusions which go beyond even the *perceived* statistical facts. Extreme and aberrant actions on behalf of the few can thus lead to conclusions concerning the group at large, and these conclusions will influence our judgments concerning a newly encountered member of the group.

A rather pristine version of this pattern of reasoning is presented 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 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.¹⁵

¹⁵ It is important to distinguish my thesis from the thesis that racist generalizations are the result of illusory correlations (see L. J. Chapman, "Illusory Correlation in Observational Report," *Journal of Verbal Learning and Verbal Behavior*, 6, (1967): 151-155; D. L. Hamilton & R. K. Gifford, "Illusory Correlation in Interpersonal Perception: A Cognitive Basis of Stereotypic Judgment", *Journal of Experimental Social Psychology*, 12 (1976): 392-407; B. Mullen & C. Johnson, "Distinctiveness Based Illusory Correlations and Stereotyping: A Meta-Analytic Integration". *British Journal of Social Psychology*, 29 (1990): 11-28.) Illusory correlations 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 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, however I believe 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

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 then we should find many examples of it, in whatever historical period we happen to examine. We should not be surprised if this mechanism of generalization has hovered perpetually in the background wherever human beings were formulating prejudiced attitudes towards social groups. A detailed historical analysis is beyond the scope of this paper, but let us briefly consider a recent and vivid example.

Nothing has done more to harm the plight of Muslims in America than 9/11. In the aftermath of 9/11, hate crimes against Muslims rose more than 1,600%, according to FBI statistics.¹⁶ 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 involved in or personally responsible for the 9/11 bombings. It was sufficient that the victims were Muslims. We might characterize the reasoning of the hate

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 Ingalls' 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".

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's experiment 2) – theorists tend to only employ them in explaining negative stereotypes, often ones involving strikingly negative behaviors of the sort here discussed. Hamilton and Gifford argue 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" (*op. cit.* 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, since it is only the perceived frequency, not the polarity or extremity that is relevant. For example, I suppose that bouncing around on a pogo stick is a very rare activity, and certainly one hears about far fewer instances of pogo-stick jumping than acts of terrorism. We would thus 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.

¹⁶ Reported in the San Francisco Chronicle,
<http://sfgate.com/cgi-bin/article.cgi?f=/c/a/2002/11/26/MN224441.DTL>

crime perpetrators as moving from the horrific events of 9/11 – events which involved a rather small number of extreme individuals – 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, now Senator, C. Saxby Chambliss told law enforcement official to “just turn [the sheriff] loose and have him arrest every Muslim that crosses the state line.”¹⁷ These statements again reflect conclusions pertaining to Muslims quite generally. They do not reflect the more moderate conclusion that only some Muslims had any involvement in 9/11 whatsoever.

Obviously the aftermath of 9/11 is not an isolated historical example. Consider, for example, the origins of anti-Algerian prejudice in France. The relationship between the French and the Algerians is a complicated one, as would be expected given their history of war, colonization and occupation, and indeed anti-Algerian racism is still virulent in France today. The cognitive bias under investigation here is, of course, too simple to account for all the subtleties of a racism with such a complex history. However, if we trace anti-Algerian prejudice to its early days, we find again the dead hand of striking property generic reasoning.

¹⁷ Reported by Human Rights Watch, <http://www.hrw.org/reports/2002/usahate/usa1102-05.htm>

Early relations between Algerian workers and France's indigenous population were quite amicable.¹⁸ 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".¹⁹

According to Neil MacMaster, this rapid swell in hostility towards North Africans can be traced to a single catalytic 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.

Immediately following this incident, there began a surge of anti-Arab violence, hatred, and discrimination throughout France. 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."²⁰ These "unconsidered generalizations" are exactly those considered in this paper. There is no *a priori* reason to think that human beings 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, however, possessed of a

¹⁸ Beaugency, 1914, as reported in N. MacMaster, "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, 1995:149-160).

¹⁹ *Ibid.*, p. 150.

²⁰ *Ibid.*, p. 158.

particular cognitive bias – a style of generalization – that makes such reasoning not only possible, but pervasive.

More speculatively, since the veil of years here is thicker, the early origins of Anglo-American prejudice towards Africans and Native Americans may have been fueled by sensational reports of horrific acts relayed in “travel logs”, which were extremely popular among the newly literate population of Britain. 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 Africans and Native Americans derived almost wholly from these travel logs.²¹ 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.²²

These travel logs, which did so much to shape England’s early image of Africa, contained endless gory accounts of shocking behavior (allegedly) exhibited by the Africans. 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

²¹ R. Cole, “Sixteenth Century Travel Books as a Source of European Attitudes toward Non-White and Non-Western Culture”, *Proceedings of the American Philosophical Society*, 116, 1 (1972): 59-67.

²² W. Jordan, *White over Black: American Attitudes toward the Negro, 1550–1812* (Chapel Hill: University of North Carolina Press, 1968), pp. 25-28.

explorers. Many of them were quite responsible in their reporting, and did not indulge themselves in broad generalizations. Given the nature of our default system of generalization, they did not have to. The reporting of specific instances would suffice to encourage very general beliefs in the mind of the reader.

Generalizations, Dispositions and Predictors

If the foregoing is correct, then the same pattern of generalization is in play for both claims like *mosquitoes carry the West Nile virus* and claims like *Muslims are terrorists*. Surely though, there must be some dissimilarities between them: in particular, is it not the case that the former claim is *true*, while the latter is *false*? Even if the same unreflective mechanism is responsible for both judgments, it is surely making an error in judging that Muslims are terrorists – an error which is not (necessarily) involved in judging that mosquitoes carry the West Nile virus.

Let us then consider striking property generics in more detail. Their truth conditions are not quite as straightforward as the earlier 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 fish attack bathers. People do not tend to find these inferences acceptable, so the truth conditions of these generics must involve some further complexity.

In earlier work, I suggested that the mechanism of generalization in question seeks a *good predictor* of the property in question.²³ 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.²⁴ 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,

²³ Leslie, 2008, *op. cit.*

²⁴ *Ibid.* 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 M. Fara, "Dispositions and habituals", *Nous*, 39 (2005): 43-82; D. Lewis, "Finkish dispositions", *Philosophical Quarterly*, 47 (1997): 143-58.

and the others are typically disposed to possess it.²⁵ To determine which striking property generics are strictly true and strictly false, then, would require some rather detailed knowledge of dispositions and capacities. Dispositions are not directly observable in the way their manifestations are, and so we do not normally possess such knowledge. We thus often operate under uncertainty when it comes to attributing dispositions, and so must adopt certain heuristics to guide our judgments. To probe this strategy 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 for 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 for most of these generalizations. However, even in the absence of scientific knowledge, we nonetheless often form (tacit) beliefs about the shared natures and dispositions of members of certain kinds. In the psychological literature, these kinds are said to be *essentialized*.²⁶ We essentialize a kind if we form the (tacit) belief

²⁵ 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 be the intuitively correct conclusion here: under such circumstances, the more inclusive generic claims are, strictly speaking, false.

²⁶ D. L. Medin & A. Ortony, “Psychological essentialism”, in S. Vosniadou & A. Ortony (eds.) *Similarity and Analogical Reasoning* (New York: Cambridge University Press, 1989: 179–196); S. A. Gelman, *The Essential Child* (New York: Oxford University Press, 2003).

that there is some hidden, non-obvious, and persistent property or underlying nature shared by members of that kind, which causally grounds their common properties and dispositions.²⁷

For example, 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.²⁸

Different levels of the subjective taxonomic hierarchy for biological kinds are essentialized to differing degrees, in the sense that some levels are seen as possessing highly distinctive essences which ground a large number of shared features, while others are not. For example, simply being a mammal is not predictive of wide range of properties, since animals as diverse as tigers, whales, mice, and humans all count as mammals. Conversely, while Bengal tigers share many properties in common with each other, this shared nature is not particularly distinctive since many of those properties are also shared by Siberian tigers. There tends to be a privileged level of the subjective taxonomic hierarchy at which the

²⁷ There is a variety of evidence to suggest that we view kinds as essentialized from a young age. For example, pre-school aged children 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 (see Gelman, *op. cit.* and references therein). They also maintain that a raccoon dressed up as 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. Children also have 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 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.

²⁸ 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. It should also 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.

essence of the kind is taken to ground a wide range of properties that are shared by its members but not collectively shared by members of another kind. This privileged level of the taxonomy is known as the *basic-level*. The notion of a basic-level kind is due to Eleanor Rosch and her colleagues, who found that various measures of psychological salience converged on a particular taxonomy that is psychologically privileged.²⁹

The most interesting feature of basic-level kinds from our point of view is that, from a young age, we see them as having rich inductive potential, thanks to their members sharing highly predictive natures. These highly essentialized basic-level kinds appear to be the default starting points for our common inductive generalizations. As members of basic-level kinds are viewed as sharing a nature, we are inclined to treat these kinds as supportive of inductive generalizations and inferences about non-obvious properties. Further, since this nature is taken to be distinctive – that is, not shared by other comparable kinds – we are reluctant to generalize properties to the more inclusive kinds above the basic-level on the taxonomic hierarchy. The psychologically privileged status of basic-level kinds is the result

²⁹ E. H. Rosch, “Principles of categorization”, in E. Rosch & B. Lloyd (eds.), *Cognition and Categorization* (Hillsdale, N.J.: Erlbaum Associates, 1978:27-48). 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.

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 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.

of a useful trade-off between the extent to which natures or essences are taken to be shared and the extent to which they are taken to be distinctive.

I suggest that, in making a striking property generalization, the default is to choose a relevant highly essentialized basic-level kind as the target of the generalization. Since such kinds are the primary targets of our inductive generalizations and inferences, it is perhaps not surprising that they are also the targets 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, 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.³⁰ I thus propose that when an instance of a particular basic-level kind manifests a striking property, 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, that is, those that share the common nature.

³⁰ 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?

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.

Social Kinds and Essence

In recent years, a number of social psychologists have argued that we view some social kinds as essentialized in much the way we view animal kinds. This line of thinking was popularized by Rothbart and Taylor,³¹ and has since received a significant amount of empirical support.³² Within psychology, the basic observation originated with Gordon Allport in 1954, who wrote that:

... a belief in essence develops. There is [a belief in] an inherent ‘Jewishness’ in every Jew. The ‘soul of the Oriental,’ ‘Negro blood,’ ... ‘the passionate Latin’ – all

³¹ M. Rothbart & M. Taylor, “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, 1992:11-36).

³² E.g.: L. A. Hirschfeld, *Race in the Making* (Cambridge, MA: MIT Press, 1996); F. J. Gil-White, “Are ethnic groups biological “species” to the human brain?” *Current Anthropology*, 42, (2001): 515-554; N. Haslam, L. Rothschild, L., & D. Ernst, “Essentialist beliefs about social categories”, *British Journal of Social Psychology*, 39, (2000): 113-127; N. Haslam, L. Rothschild, & D. Ernst, “Are essentialist beliefs associated with prejudice?” *British Journal of Social Psychology*, 41, (2002): 87–100; S. Demoulin, J. Leyens, & V. Yzerbyt, “Lay theories of essentialism,” *Group Processes and Intergroup Relations*, 9,(2006): 25-42.

A related and fascinating line of empirical research has been pioneered by Leyens and colleagues under the heading of *infra-humanization*. They present a series of empirical findings that suggest people are more reluctant to attribute uniquely human emotions to members of (at least some) out-groups. While people readily attribute to out-group members emotions that are shared with animals (e.g. fear, anger, surprise), they are less likely to attribute uniquely human sentiments (e.g. shame, resentment, love) to out-groups members vs. in-group members. The researchers argue convincingly that this reflects a tendency to deny fully human essence to certain out-groups. This work thus suggests that, not only are out-group members seen as possessing a distinctive essence, but that essence is fundamentally *less than fully human*. Conversely, one might interpret their findings as reflecting that, to the extent that the in-group is essentialized at all, its essence consists of a purely human essence. See, e.g., J-Ph. Leyens, A. Rodríguez-Pérez, R. Rodríguez-Torres, R. Gaunt, M. P. Paladino, J. Vaes, et al, “Psychological essentialism and the differential attribution of uniquely human emotions to ingroups and outgroups”, *European Journal of Social Psychology*, 31, (2001): 395–411; J-Ph. Leyens, B. Cortes, S. Demoulin, J. F. Dovidio, S. T. Fiske, R. Gaunt, et al. “Emotional Prejudice, Essentialism, and Nationalism,” *European Journal of Social Psychology*, 33, (2003): 704–717.

represent a belief in essence. A mysterious mana (for good or ill) resides in a group, all of its members partaking thereof.³³

Rothbart and Taylor argue that we may view some social kinds – such as racial or ethnic kinds – not simply as essentialized, but more specifically as natural kinds, and in some cases as biological kinds. This may be true, but it obscures the central notion. In the case of animal kinds, there is no need to separate out these different threads of belief, but in the case of social kinds it is imperative to do so. The notion of essence – or at least the notion of essence that is relevant for our purposes here – is not limited to biological kinds, or even to natural kinds (though such kinds constitute paradigmatic examples of essentialist thinking). For the remainder of this paper, I will understand a kind or group to be essentialized just in case its members are viewed as sharing a fundamental nature that causally grounds a substantial number of their outwardly observable properties. This nature need not be biologically grounded, nor need it be seen as immutable or strictly necessary for membership in the kind.³⁴ With this understanding of the notion of essence we can generalize the idea of a basic-level kind to the social arena: these will be social kinds that are perceived to have essences that occupy a ‘sweet spot’ in trade-offs between distinctiveness (which is compromised as groups become less inclusive) and predictiveness (in the sense of grounding

³³ G. Allport, *The Nature of Prejudice* (Reading, MA: Addison Wesley, 1954), pp. 173-174.

³⁴ Recent work in social psychology has begun to unravel the various threads of essence. For example, Haslam and colleagues (2000, *op. cit.*; 2002, *op. cit.*) distinguish between the perception of a group as a natural kind (e.g. having sharp boundaries, being determined by nature rather than man, being immutable) and the perception of a group as *entitative*. A group is highly entitative if its members are perceived as being very similar to one another, and if membership in the group is highly informative about the nature of the individual. Their notion of entitativity most closely corresponds to how I am here understanding essence, though I hesitate to adopt this terminology, because it is put to somewhat different uses by other social psychology researchers. (For example, Demoulin et al., *op. cit.* understand a group to be entitative if, in addition to Haslam et al.’s criteria, the group has common goals and will face a “common fate.”) Interestingly, Haslam et al. (2000, *op. cit.*) found that groups that were viewed as more entitative tended to be accorded lower status in society, but the degree to which groups were viewed as natural kinds did not predict their perceived status. B. Bastian and N. Haslam (“Psychological essentialism and stereotype endorsement” *Journal of Experimental Social Psychology*, 42, (2006): 228-235) further found that people who were inclined to essentialize social groups were more likely to endorse social stereotypes, and were also more likely to attribute the persistence of stereotypes within a culture to the nature of the group being stereotyped (i.e. as opposed to socio-cultural conditions).

the maximal number of common features – a feature which is compromised as groups becomes more inclusive). These kinds will thus be taken to have highly distinctive essences that ground a large number of shared features. Such social kinds, we may suppose, will surely include racial, ethnic, and religious kinds.

As with basic-level animal kinds, these social kinds play a privileged role in our inductive practices. The perception that they share highly predictive natures makes us willing to generalize across the group, and the perception that these natures are distinctive means that we may be reluctant to generalize to more inclusive categories – to the category *human being*, for example. As with animal kinds, it is these highly essentialized social kinds that are the typical targets of our striking property generalizations.

To recap, then, our most primitive generalizations, voiced in language as generics, are especially sensitive to information that is particularly striking, horrific, or appalling. When we encounter individuals engaging in such an act, we are naturally inclined to seek to generalize this action to a kind to which the individuals belongs. The correctness conditions of these generalizations require that some members of the kind must indeed have the relevant property, and also that the other members must be typically disposed to have the property. We do not, however, normally have good information about unobservable dispositions available to us, so as a proxy, we generalize the property to a kind that we perceive to have a highly predictive and distinctive essence. In the social domain, this means that aberrant actions on behalf of the few – such as the 9/11 terrorists, or the Native Americans involved in the Minnesota massacre, or the Algerian Khemil Ousliman, and so on – can lead us to draw conclusions about an entire social kind. To the extent that Muslims, say, are perceived as forming a deeply uniform group whose nature is quite different from other comparable

groups (e.g. Christians, Jews), conclusions such as *Muslims are terrorists* may be drawn on the basis of the actions of the few.

It is important, I think, to emphasize that this notion of essence or nature here is not a biological notion. An essence *can* of course be biologically grounded, but it need not be: a kind can be essentialized without being construed as a biological kind.³⁵ Much of the philosophical literature on race places considerable emphasis on the faulty conception of race as biologically grounded. Race is not a biological notion; there is no genetic ground for dividing people up along racial lines.³⁶ This is a tremendously important point that should be made repeatedly; however, we should not overestimate its capacity to alter people's prejudiced convictions. A belief in essence – in a shared nature – may be the more important belief to change, but beliefs in essence can survive the loss of belief in a biological essence.³⁷

³⁵ Interestingly, N. Haslam and S. Levy (“Essentialist beliefs about homosexuality: Structure and implications for prejudice”, *Personality and Social Psychology Bulletin*, 32, (2006): 471-485) found that people who endorsed biological explanations for homosexuality were less likely to evidence anti-gay prejudice. However, Haslam et al. (2002, *op. cit.*) found that the degree to which people perceived gay men to constitute an “entitative” group (see footnote 34) *did* correlate with prejudice. One potential explanation of this phenomenon is that people who believe that sexual orientation is a simple biological phenomenon do not consider gay people to have fundamentally different natures from heterosexual people, but rather view sexual orientation as a more adventitious phenomenon (more akin to differences in taste in movies, or in athletic abilities, perhaps). People who view homosexuality as a choice – and more specifically as a morally reprehensible choice – may take this decision to act immorally (by their lights) as indicative of a fundamentally different nature. (E.g. ‘I could never choose to do something so despicable, so anyone capable of making that choice must be deeply different from me and mine.) Thus belief in shared nature can come apart from beliefs about biological bases, as this case perfectly illustrates. Homosexual people – and perhaps gay men in particular (Haslam and Levy, *op. cit.*) – may still be essentialized by others even absent belief in biology. (And of course striking property generalizations concerning gays abound: *gays are pedophiles*, *gays are sexual predators*.)

³⁶ For clear and accessible discussion of the biology and its potential philosophical significance, see K. A. Appiah, “Race Culture, Identity: Misunderstood Connections”, in K. A. Appiah & A. Guttman (eds.) *Color Conscious: The Political Morality of Race* (Princeton, NJ: Princeton University Press, 1996: 30-106); K. A. Appiah, “How to Decide if Races Exist”, *Proceedings of the Aristotelian Society*, 106, 1, (2006): 365-382.

³⁷ A further concern one might have with the emphasis on the non-biological nature of race is that it seems to support the wrong counterfactuals. For example, if geneticists discovered that in fact there were significant genetic differences between races, what would follow? Using the non-biological grounds of race as the basis for discussion of prejudice – and as the cornerstone of philosophical thinking about how to combat it – seems to hold such efforts unnecessarily hostage to the deliverances of empirical science. *Greater genetic variance need not translate into deep difference in essence or nature*. The real issue here would seem to be the latter, rather than the former.

To illustrate, we may consider attitudes towards Muslims – a group that is rarely if ever taken to be *biologically* defined. The absence of belief in biological essence does not preclude belief in an essence more broadly understood, as is illustrated by persisting beliefs that Muslims are *fundamentally different* from other groups of people, while also being *fundamentally all alike*. For example, in an influential article entitled “The Roots of Muslim Rage,” emeritus Princeton professor Bernard Lewis discusses Islam and its followers. He argues that 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*” (my emphasis – note the high level of generality implied by his language).³⁸ Lewis goes on to describe the followers of Islam 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.³⁹

According to Lewis, 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 suggests 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 – and further it

³⁸ B. Lewis, “The Roots of Muslim Rage”. *The Atlantic Monthly*, 266, (September 1990).

³⁹ *Ibid.*

is one that causally grounds their putative violent dispositions. In effect, Lewis is offering a rationalization of anti-Muslim striking property generalizations.

Essence and Attribution Errors

Highly essentialized social groups are prime targets of striking property generalizations. Extreme actions committed by a very small number of the members of such a group can result in the acceptance of a generalization concerning the whole group. If the group is highly essentialized then we readily suppose that the disposition to such action is widely shared among the members – indeed that it is grounded in the very nature of the kind. When presented with a) strikingly awful actions, and b) a highly essentialized group, our primitive cognitive mechanism delivers the corresponding generalization on this basis alone.

One interesting observation is that we do not tend to make these generalizations concerning groups to which we belong. Groups of which we are members seem to be 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 our own 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 to which we belong commits an appalling act – and there is no highly essentialized sub-group that includes the individual but not us – we do not reach beyond that individual in attributing the disposition to so act. If such an individual belongs to a highly essentialized group that excludes us, 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 thereby identified another 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.⁴⁰ There has been some discussion of another attribution error, known as the Ultimate Attribution Error (UAE), which is concerned with attributions across groups.⁴¹ 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.⁴²

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 a way that parallels our tendency to explain away our own bad behavior as situationally determined while attributing the negative actions

⁴⁰ E. E. Jones & R. E. Nisbett, “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, 1972: 79-94).

⁴¹ T. Pettigrew, “The Ultimate Attribution Error: Extending Allport’s Cognitive Analysis of Prejudice”, *Personality and Social Psychology Bulletin*, 5, 4, (1979): 461-476.

⁴² For a review, see M. Hewstone, “The ‘Ultimate Attribution Error’? A Review of the Literature on Intergroup Causal Attribution”, *European Journal of Social Psychology*, 20, (1990): 311-335. Some of the best evidence in support of the Ultimate Attribution Error may be found in work done on the language used to describe actions. There is a tendency to describe negative actions committed by in-group members in more situational, less characterizing language than that used to describe negative actions committed by out-group members. See A. Maass, D. Salvi, L. Arcuri, & G. Semin, “Language use in intergroup contexts: The linguistic intergroup bias”, *Journal of Personality and Social Psychology*, 57, (1989): 981-993.

of out-group members to more persistent “inner” factors. The Ultimate Attribution Error thus offers the same broad contrast as the Fundamental Attribution Error. However, 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.⁴³

To illustrate the difference between the UAE and the attribution error I am discussing here, consider the reactions of Americans to the Oklahoma City bombings, and to the bombings of 9/11. To commit the UAE here, one would have to attribute Timothy McVeigh’s actions to situational factors, while refraining from offering such excuses for the 9/11 bombers. I imagine that few people are inclined to think that McVeigh was simply a victim of circumstances. By contrast, to commit the attribution error that I am describing, one can believe 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 – in

⁴³ To the best of my knowledge, this particular attribution error has not been tested, nor even discussed. It is a wholly empirical question whether we are in fact disposed to commit this attribution error – my aim here is to put it forward as a testable hypothesis. The closest empirical test that I have been able to uncover in the literature is V. Y. Yzerbyt, A. Rogier, & S. Fiske, “Group Entitativity and Social Attribution: On translating situational constraints into stereotypes,” *Personality and Social Psychology Bulletin*, 24, (1998): 1090–1104. They found that for more “entitative” groups (see footnote 34), we tend to overestimate the extent to which behavior of members is attributable to the groups’ nature vs. to situational characteristics. When the groups in question were not perceived as entitative, this effect was not found. However, the behavior in question here was mildly positive rather than strongly negative, and importantly no asymmetry was found depending on whether the group being assessed was an in-group or an out-group. It is precisely in this respect that I would predict an asymmetry in the case of strikingly awful behavior, though not in the case of bland or positive behavior.

the words of Bernard Lewis – an ‘explosive mixture of rage and hatred’ inherent in Muslims.⁴⁴

It bears emphasizing once again that similar conclusions are not drawn when we are faced with neutral or positive information. This bias is not traceable to the essentialization of these kinds alone. If we look out of our window one morning and see a group of Muslims jumping on pogo sticks, we are not inclined to conclude something very general about Muslims. We do not suppose that the essence of the group grounds the disposition to jump on pogo sticks. When we encounter Muslim doctors, we do not conclude *Muslims are doctors*, and suppose that the non-doctors among them are nursing a yet-unmanifested disposition to practice medicine. And so on with an open-ended range of examples. The disposition to generalize in this way is specific to the strikingly negative, to the threatening, the vile, the horrific, and the dangerous.

⁴⁴ This attribution bias should not be confused with the out-group homogeneity bias (G. A. Quattrone and E. E. Jones, “The Perception of Variability Within Ingroups and Outgroups: Implications for the Law of Small Numbers”, *Journal of Personality and Social Psychology*, 38, (1980): 141-152; E. E. Jones, G. C. Wood & G. A. Quattrone, “Perceived Variability of Personal Characteristics in In-groups and Out-groups: The Role of Knowledge and Evaluation”, *Personality and Social Psychology Bulletin*, 7, (1981): 523-528; B. Park & M. Rothbart, “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*, 42, (1982): 1051-1068).

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 (B. Mullen & L. Hu, “Perceptions of Ingroup and Outgroup Variability: A Meta-Analytic Intergration”, *Basic and Applied Social Psychology*, 10, 3, (1989): 233-252). 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 (A. Guinote, “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*, 40, (2001): 117-132; B. Simon & R. Brown, “Perceived Intra-group Homogeneity in Minority-Majority Contexts” *Journal of Personality and Social Psychology*, 53, (1987): 703-711; B. Simon & T. Pettigrew, “Social Identity and Perceived Group Homogeneity: Evidence for the Ingroup Homogeneity Effect”, *European Journal of Social Psychology*, 20, (1990): 269-286). It would be surprising, though, if minority members were inclined to make these striking property generalizations over their own groups, however (though the question is ultimately an empirical one, of course). 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.

Looking for Hope: Familiarity and Identification

That we resist making such generalizations concerning groups to which we belong does not, on the face of it, hold much promise for helping to reduce the tendency to generalize in this way. Simply put, one cannot be a member of every essentialized group. However, increasing familiarity, knowledge, and solidarity may serve as a surrogate of sorts: people who are very familiar or identified with a class seem to reject striking property generalizations over that class. Those with considerable experience of 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, yet few if any of us are inclined to accept the generalization *fathers molest children*. A key factor seems to be a significant degree of *familiarity* or even *solidarity* with members of the relevant kind.

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.

If the search for less inclusive groups were carried to its ultimate conclusion, we would simply treat each individual wholly on their own terms. (As discussed above, if there are no available essentialized kinds that include the perpetrator but not ourselves, this may be exactly what we do.) Even short of this, however, there is much room for improvement. For

example, we might prefer such generalizations as *disturbed fanatical extremists blow up buildings* in place of *Muslims blow up buildings*.

This search for increasing specificity may be part of a more general tendency, as familiarity and knowledge grows, to view increasingly restrictive sub-kinds as comprising the basic-level. In the biological domain, for example, recent findings suggest that the Itzaj Mayan people treat individual tree species as highly essentialized basic-level kinds, and the category *tree* as a less essentialized superordinate kind. American college students, however, treat *tree* as a basic-level kind. A natural explanation is that the Itzaj Mayans are simply more knowledgeable about trees than the American college students, and so the differences between the sub-kinds of trees are too salient to them to be ignored.⁴⁵ (Consider the neophyte wine drinker to whom there seems to be just two kinds of wine: red and white. In contrast, the most discerning oenophile might see each triple of vineyard, grapes and vintage as carving out so unique a kind that each must be considered solely on its own terms, with no higher generalizations permitted.)

If the above observations concerning familiarity, knowledge and solidarity are true, this would further serve to expose the fatal, reinforcing “logic” of segregation. The more separated and distant we are, the more we see each other in the terms that invite invidious striking property generalizations.

These claims fit nicely with the so-called contact hypothesis, originally advanced by Gordon Allport in 1954.⁴⁶ He predicted that contact between members of 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

⁴⁵ J. D. Coley, D. L. Medin, J. B. Proffitt, E. Lynch, & S. Atran, “Inductive Reasoning in Folkbiological Thought”, in D.L. Medin & S. Atran (Eds.), *Folkbiology* (Cambridge, MA: The MIT Press, 1999: 205-232).

⁴⁶ *Op. cit.*

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.⁴⁷ 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 lead us to resist viewing broadly inclusive social kinds as highly essentialized.

Looking for Hope: Language and Essentialism

Familiarity and knowledge can get us to view increasingly less inclusive kinds as the basic-level kinds. We might also consider which factors originally lead us to view some social kinds as possessed of widely shared and highly distinctive natures, and whether anything can be done to undermine this tendency. When such a belief in the group's shared nature is lacking, we do not tend to accept these generalizations – for example, we do not think that *accountants are murderers* even though some accountants have been known to commit murder, nor do we think that *fish attack bathers* even though some fish – that is, some sharks – do indeed attack bathers. A belief that such a shared essence exists and grounds the disposition to the striking behavior is a precondition for these generalizations.

If I am correct to suppose that striking property generalizations have played a role in the formation of some of our social prejudices, then the question arises, how might we most effectively combat and undermine these generalizations? The claim that the essentialization of these groups is a precondition here raises a tantalizing empirical suggestion: if we can

⁴⁷ T. Pettigrew & L. Tropp, “Does Intergroup Contact Reduce Prejudice? Recent Meta-Analytic Findings”, in S. Oskamp (ed.) *Reducing Prejudice and Discrimination* (Mahwah, N.J.: Lawrence Erlbaum Associates, 2000: 93-115).

identify factors which lead us to essentialize a given kind, then we may be able to develop novel means of combating prejudice.

Recent work by Susan Gelman and her collaborators is, I think, particularly telling. Their focus has been primarily on the development of essentialist thinking about animals and other biological kinds. By age 4, children have essentialist beliefs about such kinds.⁴⁸ Since it seems that we may essentialize social kinds a similar manner, this invites the hypothesis that factors lead us to essentialize biological kinds might perhaps also contribute to our essentialization of non-biological social kinds.

It is important to notice that, while the disposition to essentialize *some kinds or others* may be an immutable feature of our cognition, this does not determine *which* kinds we essentialize. While the available cross-cultural evidence suggests that essentialization may be a universal human phenomenon, there is significant cross-cultural variation in *which* kinds are seen as essentialized.⁴⁹ For example, the different castes in India have been seen as highly essentialized, especially by members of the upper castes.⁵⁰ The Ancient Greeks are often said to have believed that there were two fundamental kinds of human beings: Greeks and Barbarians, each endowed with their own distinctive natures. We might suppose that class in English society has, at least until very recently, been essentialized, and certainly medieval European notions of the Great Chain of Being involve highly essentialist ways of

⁴⁸ Gelman, 2003, *op. cit.*

⁴⁹ E.g., S. Atran, "Itzaj Maya folkbiological taxonomy: Cognitive universals and cultural particulars", in D.L. Medin & S. Atran (Eds.), *Folkbiology* (Cambridge, MA: The MIT Press, 1999: 233-284); Coley et al., *op. cit.*; Gil-White, *op. cit.*; R. Astuti, G. E. A. Solomon & S. Carey, "Constraints on conceptual development." *Monographs of the Society for Research in Child Development*, 69, 3, 277, (2004); for a review, see G. Diesendruck, "Categories for names or names for categories? The interplay between domain-specific conceptual structure and language", *Language and Cognitive Processes*, 18, (2003): 759-787.

⁵⁰ R. Mahalingam, *Essentialism, Power, and Representation of Caste: A Developmental Study* (Ph.D. Dissertation, University of Pittsburgh, 1998).

thinking about the different strata of society. (Rebellious peasants were often said to be going against *nature itself*.) It is not difficult to come up with many more such examples.

Thus, while the capacity to essentialize may be an innate feature of human psychology, it does not fix *which* kinds are to be essentialized. We are not hard-wired to form the specific social categories that we do.⁵¹ The examples just noted do not involve slicing humanity along the same lineaments as contemporary American thinking on race or ethnicity: the Greek term ‘barbarian’ encompassed a wide variety of groups which we would distinguish in contemporary American society, while the social classes demarcated in medieval Europe cut more finely than the racial and ethnic categories which are so salient to us.

Since it is not a fixed and immutable part of our psychology to essentialize the specific groups that we do, the urgent question arises, what determines which groups we essentialize? As with everything else in this domain, I am highly skeptical of the possibility of providing a single, univocal answer to this question. The issue here as elsewhere is partly historical, partly economic, partly social, and partly who-knows-what-else. There is a variant on the question, though, which may prove more tractable from the cognitive perspective. I propose that we instead ask the question: what are some of the cues that suggest to young children that a particular kind or group is to be essentialized?

⁵¹ For a fascinating empirical demonstration of how malleable our social categories (other than gender) may be, see R. Kurzban, J. Tooby & L. Cosmides, “Can race be erased?: Coalitional computation and social categorization”, *Proceedings of the National Academy of Sciences*, 98, 26, (2001): 15387-15392. These authors argue that, contra prior claims, racial categories are not automatically encoded when we perceive others. Rather, we primarily encode information about membership in coalitions. The authors argue that the encoding of race is just a special case of this more general tendency – one that gets reinforced over a lifetime’s participation in our society, thus acquiring an air of inevitability. It can be significantly undermined in certain contexts, however.

Though the question is far from settled, recent work on children's understanding of race suggests that racial groups come to be essentialized only later in development.⁵² This accords well with the observation that there is significant cross-cultural and historical variation in which social kinds are essentialized – we do not come into the world hard-wired to essentialize racial kinds, and presumably the same also holds for ethnic and religious kinds.⁵³ This is something that we have to *learn* to do.

How, then, do children learn to essentialize one kind rather than another? This question is beginning to be investigated, especially in the case of animal kinds. If the same cognitive processes are implicated in the essentialization of both animal and social kinds, then we might form the tentative hypothesis that findings concerning the one will apply to the latter. Current research is far from providing a full account of the phenomenon; however, it does point to some factors which can increase the extent to which children and adults essentialize particular kinds; in particular, recent findings suggest that the *language* we use to talk about individuals and groups can have a significant impact on the degree to which they are essentialized.⁵⁴ Here, as elsewhere in psychology, something that looks on its face to be

⁵² For example, M. Rhodes & S. A. Gelman, "A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts", *Cognitive Psychology*, in press. Interestingly, gender may be essentialized significantly earlier (e.g. Gelman, 2003, *op. cit.*; M. G. Taylor, M. Rhodes, & S. A. Gelman, "Boys will be boys, cows will be cows: Children's essentialist reasoning about human gender and animal development", *Child Development*, 79, (2009): 1270-1287). And unlike racial categories, evidence suggests that gender categories are strongly and automatically encoded in person perception (Kurzban et al., *op. cit.*); see footnote 51.

⁵³ For an evolutionary perspective on this point, see L. Cosmides, J. Tooby & R. Kurzban, "Perceptions of Race", *Trends in Cognitive Sciences*, 7, (2003): 173-179; also Kurzban et al., *op. cit.*

⁵⁴ On this point, I am in agreement with Hirschfeld, 1996, *op. cit.*; D. Sperber, *Explaining culture: A naturalistic approach* (Cambridge, MA: Blackwell, 1996); and S. Carey, "On the origin of causal understanding," in D. Sperber, D. Premack, & A.J. Premack (eds.), *Causal cognition: A multi-disciplinary debate* (Oxford: Oxford University Press, 1995: 268–308. Subsequent empirical findings have, I would argue, supported these theorists' insights in this regard. See also Diesendruck, 2003, *op. cit.*

superficial may have a deeper significance because of the idiosyncratic character of our species' cognitive structure.⁵⁵

One recent study is quite telling on this point. Susan Gelman, Elizabeth Ware, and Felicia Kleinberg designed three picture books about a novel animal, which differed from each other only in the language they used to talk about the animals depicted therein.⁵⁶ The first type of book made frequent use of generic language (e.g. "zarpies hate ice cream"), while the second contained no generics, but did make use of the common noun 'zarpie' (e.g. 'this zarpie hates ice cream'). The third book did not contain the common noun 'zarpie' at all, and the individuals were simply referred to as 'this' (e.g. 'this hates ice cream'). A group of parents agreed to read one of these picture books frequently to their four year old children over a period of several weeks. The children and their parents were tested (separately) with a battery of tests designed to assess the extent to which they were prepared to treat zarpies as an essentialized kind (e.g. the extent to which they viewed zarpie traits as innate vs. learned, the extent to which they generalized novel properties from one zarpie to another, and a number of other such measures).

Children and adults who saw/heard the zarpies described with generic noun phrases showed by far the greatest tendency to essentialize zarpies. This effect held up over a period of several weeks, including weeks in which the children were no longer actively reading the picture books, which suggests the effect may be quite robust. (In some respects, the effect of

⁵⁵ In what follows, I wish to be clear that I am not endorsing the so-called Whorfian view that language *determines* cognition in any non-obvious way. Like many of the researchers on whose work I draw (e.g. Susan Gelman) I believe that the results and implications discussed can be readily accommodated on the rather bland supposition that language conveys information, both semantically and pragmatically, and that the information we grasp has an obviously important impact on how we conceptualize the world. For example, a child that hears a kind repeatedly described with generic language may come to essentialize that kind *because the child recognizes that the surrounding adults see this as a kind worthy of discussion in generic terms*. This recognition need not be conscious in order to be explicable in pragmatic terms.

⁵⁶ S. A. Gelman, E. A. Ware & F. Kleinberg, "Effects of Generic Language on Category Content and Structure", *under review*.

hearing generics actually grew *more* pronounced as time went on.) This study thus offers clear evidence that language plays a role in shaping the extent to which we essentialize a given kind.⁵⁷

Other studies have compared the use of nouns/labels to verbs or adjectives/descriptions, and found that the former foster essentialism, or at least related phenomena.⁵⁸ For example, Susan Gelman and Gail Heyman⁵⁹ found that five and nine year olds drew different conclusions depending on whether they were told, e.g., “Rosie eats carrots whenever she can” vs. “Rosie is a carrot-eater”. The use of the common noun led children to view the property as more stable, and they judged that Rosie was more likely to persist in eating carrots in the absence of the usual parental encouragement.⁶⁰ In an induction task conducted with Israeli preschoolers, Gil Diesendruck and Heidi haLevi found that five-year-olds preferred to generalize novel properties based on membership in familiar social categories as conveyed by labels (e.g. a Jew vs. an Arab), rather than based on either shared

⁵⁷ Zarpies, as presented in the study, are intended to comprise a novel *animal* kind. Will the same effects of language be found if the group in question is a *social* kind? Marjorie Rhodes and I are currently collaborating to test this hypothesis – there is no data available at time of writing, however.

⁵⁸ For more discussion of nouns vs. adjectives, see A. Wierzbicka, “What’s in a noun? (Or: How do nouns differ in meaning from adjectives?)”, *Studies in Language*, 10, (1986): 353–389; see also Carey, 1995, *op. cit.*

⁵⁹ S. A. Gelman & G. D. Heyman, “Carrot-eaters and Creature-believers: The effects of lexicalization on children’s inferences about social categories”, *Psychological Science*, 10, (1999): 489–493.

⁶⁰ Walton and Banaji found that the same held for adults, and Markman and Smith report a related contrast between nouns and adjectives (G. M. Walton & M. B. Banaji, “Being what you say: The effect of essentialist linguistic labels on preferences”, *Social Cognition*, 22, (2004): 193–213; Markman and Smith’s work is reported in E. M. Markman, *Categorization and naming in children: Problems of induction* (Cambridge, MA: MIT Press, 1989)).

In a fascinating recent study, Carnaghi, Maass, Gresta, Bianchi, Cadinu, and Arcuri (“*Nomina Sunt Omnes*: On the Inductive Potential of Nouns and Adjectives in Person Perception”, *Journal of Personality and Social Psychology*, 94, 5, (2008): 839-859) extend a similar paradigm to social kinds, and again find a parallel contrast between nouns and adjectives in the responses of their adult participants. These studies on adults tend to focus on the elicitation of previously held beliefs (with the exception of Walton and Banaji’s experiment 1), however, rather than on the formation of novel ones. Since we are primarily concerned with formation rather than elicitation, my focus is on the studies that investigate the impact of language on the formation of new beliefs about essentialism.

physical traits or shared personality traits as conveyed by appearance and/or description (e.g. shy vs. friendly).⁶¹

It must be emphasized that these linguistic cues are not *sufficient* for a kind's being essentialized.⁶² We can and do use common nouns and generics to talk about non-essentialized kinds. In Gelman, Ware, and Kleiberg's experiment, it is important to notice that zarpies are clearly depicted as a type of biological creature – that is, language aside, the depicted zarpies formed a kind that was ripe for essentialization. (An interesting control task would be to look at the effects of language on less essentializable kinds, such as artifacts.) We should therefore not be misled into supposing that these linguistic cues are in and of themselves *sufficient* for essentialization of the kind. A very interesting question, though, is the extent to which these cues are *necessary* (or at least close to necessary – that is, only rarely would these cues not be needed). Certainly, it is difficult if not impossible to think of an essentialized kind for which we do not have a labeling common noun. Even if such cues are not even strictly necessary, the evidence points to their being at the very least important contributing causes. One might thus reasonably hypothesize that altering these cues would have a powerful effect.

The above results and others point, I think, to a novel proposal for combating prejudice. If prejudice and essentialism about social kinds are crucially linked, then we might suppose that undermining the latter will help mitigate the former. No studies to date have looked at possible ways to *reduce* existing essentialist beliefs, but the ones described above

⁶¹ G. Diesendruck & H. haLevi, "The role of language, appearance, and culture in children's social category based induction", *Child Development*, 77, (2006): 539-553.

⁶² Carey, 1995, *op. cit.* seems to suggest that labeling may be both necessary and sufficient for essentialism: "Essentialism, like taxonomic structure, derives from the logical work done by nouns ... I submit that biological essentialism is the theoretical elaboration of the logical-linguistic concept, substance sortal" (p. 267-277). Arguably, however, she has in mind a rather different conception of essentialism than the one under discussion here.

have identified some factors which promote essentialist thinking about a given kind. In particular, they shed some light on the cues that young children use to determine *which* kinds they should essentialize. The empirical evidence suggests that the use of nouns as labels promotes essentialism, and that the use of generics compounds this.

The intriguing upshot of this is that our very choice of words to describe racial/ethnic/religious kinds may subtly communicate to our children that these kinds are to be essentialized. We need not say anything negative about these groups – the use of generics or even simply labeling nouns may communicate that these are essentializable groups, and so open the door to prejudice.

The contrast between nouns on the one hand and verbs and adjectives on the other suggests a possible way of mitigating this effect. In the early days of research on autism, researchers would often speak of “autists” or “autistics” – using a noun to label this group of people. It was then thought that this promoted a certain undesirable way of thinking about this group, and so researchers were urged to instead speak of “autistic people” – using an adjective instead. However, this sort of adjective-noun compound is all too easily heard as just another common noun unit (compare “Bengal tiger”). Nowadays, the preferred locution is instead “people with autism” – a locution which emphasizes that they are people first and foremost, and that autism is just one property among many which they possess. The condition does not define them.⁶³

The role of common nouns – as opposed to other linguistic devices such as adjectives – in guiding categorization and generalization has been extensively studied in the non-social arena. A range of data suggest that common nouns have an impact on how we categorize and

⁶³ For empirical confirmation of the effects of such rephrasing, see C. J. Cunningham, *Illnesses as Labels: The Influence of Linguistic Form Class*, (Undergraduate Honors Thesis, University of Michigan, 1999).

generalize from as young as fourteen months of age.⁶⁴ Interestingly enough, these effects are only found when the noun is clearly presented in a ‘naming phrase’ (e.g. “A blicket!” but not “Blicket!”), reflecting that the common noun must be clearly used to indicate that the item in question belongs to a *kind*.⁶⁵ Further, for infants at least as young as sixteen months, using a description or an adjective (e.g. “this is blickish!”) in place of a noun does not produce the same effect.⁶⁶ These infants were significantly less likely to generalize a property across perceptually dissimilar items introduced by the same description, as opposed to dissimilar items introduced with the same common noun labels.

Recent and innovative work by Sandra Waxman⁶⁷ has begun to extend these findings into the social realm. Waxman presented preschoolers with pictures of people of different races, and told them that one of the people depicted had a novel property (e.g. “like to play a game called zaggit”), then asked the children whether they thought the other people depicted would also have the property in question. In the first study, no labels were used, and Waxman found that the preschoolers did not use race a guide to their attributions – rather

⁶⁴ C. Fennell & S. R. Waxman, “What paradox? Referential cues allow for infant use of phonetic detail in word learning”, *Child Development*, in press; S. R. Waxman, “Specifying the scope of 13-month-olds’ expectations for novel words”, *Cognition*, 70, (1999):B35-B50; S. R. Waxman & A. E. Booth, “Seeing pink elephants: Fourteen-month-olds’ interpretations of novel nouns and adjectives”, *Cognitive Psychology*, 43, 3, (2001): 217-242; S. R. Waxman & A. E. Booth, “The origins and evolution of links between word learning and conceptual organization: New evidence from 11-month-olds.” *Developmental Science*, 6, 2, (2003): 130-137.

⁶⁵ Fennell and Waxman, in press, *op. cit.*

⁶⁶ J. Keates & S. A. Graham, “Category labels or attributes: Why do labels guide infants’ inductive inferences?”, *Psychological Science*, 19, (2008): 1287-1293.

⁶⁷ S. R. Waxman, “Names will never hurt me? Naming and the Development of Racial and Gender Categories in Preschool Aged Children”, *European Journal of Social Psychology*, in press.

As best as I am able to tell, until recently very little work had been done on the social implications of using common nouns and/or generics. In a seminal paper, Semin and Fiedler characterized the differential social impact of some types of language, and argued that the more abstract the language, the more potent the social impact. Adjectives (e.g. “John is aggressive”), argue Semin and Fiedler, are more abstract and therefore more socially potent than, say, descriptions of actions (e.g. “John hit Mary”). However, Semin and Fiedler do not consider any more abstract forms of language than adjectival descriptions. I would argue that common noun labels are more abstract (in their sense) and thus more potent than adjectives, and further than generics are more abstract and potent forms of speech still. See G. Semin & K. Fiedler, “The Cognitive Functions of Linguistic Categories in Describing Persons: Social Cognition and Language”, *Journal of Personality and Social Psychology*, 54, (1988): 558-568.

they attributed the property just as readily across racial boundaries as they did within. In her second study, however, the person to whom the novel property was attributed was either introduced with a description (e.g. “this one likes to eat big lunches”) or with an unfamiliar common noun (e.g. “this one is a Wayshan”). Waxman found that the children who heard the description were again just as likely to attribute the novel property across racial boundaries as within, but this was not so for the children who heard the common noun. These children tended to not generalize the novel property across racial boundaries, but rather confined their generalizations to members of the same racial group. Of course these were of necessity rather artificial conditions, but this ingenious experiment clearly highlights the impact of labeling – as opposed to merely describing – on preschoolers’ social thinking.⁶⁸

Another dramatic illustration of the power of labeling in social situations is due to Zijong He and Renee Baillargeon.⁶⁹ In their experiment, 35-month-olds watched three adults interact with each other. The three adults introduced themselves using unfamiliar labels: the first adult said “I’m a mayloo”, the second said “I’m a boga”, and the third said “I’m a boga, too”. In one condition, the first adult then left, leaving the two bogas together. The second boga made a friendly gesture towards the first boga, and the gesture was either welcomed or rebuffed. The children were surprised when the gesture was rebuffed in this condition (as indicated by a looking-time measure). In another condition, the second adult was a mayloo rather than a boga. The same interaction ensued (the first mayloo left, and then the boga made a friendly overture toward the mayloo), but here the children were not surprised when the gesture was rebuffed – in fact, they looked longer when the overture was reciprocated. These results indicate that as young as 35 months, toddlers have expectations about intra-

⁶⁸ Waxman also conducted the same test for gender categories with similar results.

⁶⁹ Z. He and R. Baillargeon, “Early Expectations about In-group and Out-group Agents”, *in preparation*.

versus inter-group interactions, and that these expectations can be triggered by no more than a difference in common nouns.⁷⁰

The foregoing discussion suggests that we might consider altering our ways of speaking about race, ethnicity, religion, and so on. Instead of *labeling* a person as *a Muslim*, we might instead *describe* the person – if needed – as, say, *a person who follows Islam*, thus emphasizing that *person* is the relevant kind sortal, and that *following Islam* is a particular property that the individual happens to possess. Instead of speaking of *Blacks* and *African Americans* we might instead adopt locutions such as *people with darker skin*.⁷¹ (We might notice that already such a locution undermines the viability of essentialist thinking about race – how much sense does it make to, e.g., persist in looking for genetic differences and inherent IQ differences between racial groups, if we think through the lens of this locution?)

It might also be advisable to take some care in choosing our adjectival phrases. In particular, we should take care to avoid phrases which – in their own way – evoke an essentialist way of thinking. An extreme example of this would be, e.g., *people who share a genetically determined nature with other darker-skinned people*, but there are less dramatic possibilities which nonetheless may still promote essentialism. I know of no empirical data on this, but we might conjecture that reference to ancestry, blood, nature, and other related concepts would make for less-than-ideal adjectival phrases. Thus we might prefer *people*

⁷⁰ At time of writing, the experimenters have not explored the contrast between nouns and adjectives using this paradigm; this would be a fascinating further study however.

⁷¹ The locution *people of color* would also seem to be a step in the right direction. We might, however, prefer an even more obviously descriptive locution, since there is perhaps still something conducive to essentialization in the mysterious “of color”.

Of course *people of color* is most commonly used as a general term for non-White people rather than for Black people specifically. While such generality is not a necessary feature of a descriptive locution, it does highlight an interesting issue: could any genuinely descriptive phrase be strictly extensionally equivalent to a common noun label for an essentialized kind? I think the answer here is ‘no’ – if we take a kind to be essentialized, we expect the common noun label to apply just in case an individual partakes of the relevant essence, even if the individual does not superficially fit the descriptive bill for category membership. (Such a mindset is, of course, a pre-requisite for making sense of notions like “passing” – e.g., looking White despite *really* belonging to a non-White category.)

with darker skin – with its focus on outward, superficial properties – to *people of African descent*, since the latter leaves more conceptual room for thinking in terms of shared, hidden, heritable properties.

Adopting such a way of speaking and thinking may have some immediate benefits; the results of Carnaghi and colleagues⁷² suggest that hearing a member of a familiar social kind described by an adjective rather than a noun can reduce the extent to which adults expect the individual to conform to a stereotype. It is possible, though, that the real benefits would not lie just in the alteration of the attitudes of adults; the really intriguing possibility would be to decrease the extent to which children in our society grow up essentializing social kinds. The empirical results summarized here do not, of course, account for all the factors that lead children to essentialize, nor do they purport to. However, the evidence strongly suggests that the use of labels and generics contribute to essentialization, and so we may expect that the converse will also hold: reducing the use of labels and generics for racial, ethnic, and religious groups may reduce the extent to which children grow up essentializing these groups.

Conclusion

Depending on one's point of view, this linguistic revisionism may seem more or less appealing. The philosophical literature on race very often centers on the question of whether racial categories a) *can* be said to exist, and b) *ought* to be said to exist. Theorists are divided on this latter issue because of a number of complex reasons having to do with broad social, cultural, political, and economic factors.⁷³ I will not venture a view on these issues here,

⁷² Carnaghi et al., 2008, *op. cit.*

⁷³ Some excellent examples encompassing a variety of viewpoints include, K. A. Appiah, "The Uncompleted Argument: DuBois and the Illusion of Race", in L. A. Bell and D. Blumenfeld (eds.) *Overcoming Racism and Sexism* (Lanham, MD: Rowman & Littlefield, 1995: 79-79); K. A. Appiah, 1996, *op. cit.*; S. Haslanger,

though of course my proposal will be far more appealing to theorists like Anthony Appiah and Naomi Zack, who believe that we should eschew racial categories for political, social, and ethical reasons.⁷⁴ This work adds one consideration in favor of the stance they advocate: we may be wired in a way that makes having essentialized social kinds without concomitant prejudice extremely difficult. Eschewing reference to social kinds, at least by way of common nouns and generics, may be a direct way to reduce the extent to which we unconsciously teach our children to essentialize.

By comparison with Appiah and Zack, my perspective in this paper has been considerably more narrow – I have explored these issues solely from the point of view of cognitive psychology. The best remedy for racial injustice from the political point of view may not be the same as, or even wholly consistent with, the best preventative strategy from the cognitive point of view. To raise just one difficult question: can the members of the aggrieved group operate effectively in the political and cultural arenas without their publicly essentializing the group in *positive* ways, thereby paradoxically taking on the risk of reinforcing the very kind of prejudice I have described?

The perspective offered here is not intended as the final word on the matter – nothing could be further from my intent. My aim has rather been to take a small slice of a huge phenomenon, and identify some of the contributing cognitive factors. Identifying these

"Gender and Race: (What) Are They? (What) Do We Want Them To Be?" *Noûs*, 34, 1, (2000): 31-55; C. Mills, *Blackness Visible: Essays on Philosophy and Race* (Ithaca, NY: Cornell, 1998); L. Outlaw, "On W.E.B. DuBois's 'The Conservation of the Races'", in Bell and Blumenfeld, 1995, *op. cit.*, 79-102; L. Outlaw, *On Race and Philosophy* (New York: Routledge, 1996); R. Sundstrom, "Racial Nominalism" *Journal of Social Philosophy*, 33, 2, (2002): 193-210; P. Taylor, "Appiah's Uncompleted Argument: Du Bois and the Reality of Race", *Social Theory and Practice*, 26, 1 (2000): 103-28; N. Zack, *Race and Mixed Race* (Philadelphia, PA: Temple University Press, 1993); N. Zack, *Philosophy of Science and Race*. (New York: Routledge, 2002).

For a wonderfully illuminating and comprehensive discussion of how some socio-political considerations may interact with cognitive considerations, see D. Kelly, E. Machery & R. Mallon, "Racial Cognition and Normative Racial Theory," in J. Doris, S. Nichols and S. Stich (eds.) *The Oxford Handbook of Moral Psychology* (Oxford: Oxford University Press, forthcoming).

⁷⁴ K. A. Appiah, 1995, *op. cit.*; K. A. Appiah, 1996, *op. cit.*; N. Zack, 1993, *op. cit.*; N. Zack, 2002, *op. cit.*

contributing cognitive factors leads us to pose the question of how best to alter those factors, and thereby combat those aspects of prejudice that arise from “the original sin” of cognition, namely its primitive tendency to generalize striking negative information across the members of highly essentialized kinds.