An Examination Of The Extent Of
Adults’ Social Essentialism

by Robert Cole Eidson

B.S. in Psychology, University of Louisville
M.A. in Psychology, Northeastern University

A dissertation submitted to

The Faculty of
the College of Science of
Northeastern University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

August 5, 2014

Dissertation directed by

John Coley
Associate Professor of Psychology
Acknowledgements

While a graduate student produces an original body of work for his or her dissertation in the pursuit of their doctoral degree, it is impossible to fathom that such a feat could be undertaken alone. This endeavor is supported conceptually and physically by those who help refine ideas and appropriate research questions and help in the day-to-day collection of data as well as emotionally and psychologically by the support network that all of us requires to survive such an undertaking.

First and foremost, I must thank my advisor, Prof. John Coley for his guidance in helping to produce my dissertation. Although we all enter graduate school with the ambition and drive to become independent scientists, these are not enough to produce an original body of work over the course of several years. Instead, our advisors and mentors help to shape the scientists we will become. When I initially interviewed for graduate school, John was very clear that I would have the academic freedom to pursue questions and ideas that I found interesting helping to prepare me to ask new and exciting research questions from my own point of view. That promise became very real when I embarked on a separate line of research in our laboratory by examining people’s beliefs and intuitions about social groups. This line of work has spawned numerous projects but all with the support of Prof. Coley. Although we may sometimes be too much alike for our own good, I am honored to call him a colleague and a friend.

While I may have originated many new research questions in our laboratory and for the purposes of my dissertation, these would have been impossible to answer were it not for other graduate students, lab managers, and myriad research assistants who have come through our lab. Of course, such a list must include Categorization and Reasoning Lab alumnus Nadya Vasilyeva. When I first entered graduate school, Nadya was approaching the end of her dissertation, but she
always had time to answer my questions and provide invaluable insight into the daily research challenges that I have faced. I have been lucky to retain her input to help me refine my thinking around the ins and outs of my work. Second, these research endeavors would have come to a grinding halt if not for the lab research coordinators that have held co-op positions with us throughout my graduate career: Amanda Civiletto, Mary Angilly, Stephanie Silveira, Christine Torreson, Alison Robey, Tiffany Parkhurst, Lindsay Harrigan, John McKenna, Melanie Arenson, Sam Daoust, Lindsay Houck, and Grace Kreher. While the research assistants who have helped with data collection and analysis, experimental design, poster presentations, etc. are too numerous to mention, I sincerely thank them for their work on these projects. Specifically, I would like to thank a previous CRL research assistant whom I never had the pleasure of meeting – Marina Fontaine – for the images used in Experiment 5.

Of course, as important as the day-to-day mechanics of research are in helping to produce a dissertation, the support network one must rely on through both the ups and downs of graduate school is equally if not more important. At the center of that network are my parents and family. They have supported me in the pursuit of my interests even when they had much rather me pursue different paths. Yet, their unwavering support has made it possible to persevere through difficult times. More broadly, my lifelong friends – Hannah Wohltjen, Tyler Graham, Jeremy Snyder, Josh Hall, David Scholla, Cody Pruitt, and others – have provided me with a sounding board whenever I needed reassurance to get through the day. While those friends have provided support from a distance, the graduate students at Northeastern have also been a wonderful asset either in helping to formulate new ideas or simply providing a wonderful and friendly atmosphere in which to work and grow. Specifically, I have to recognize several current and former members of the Language and Cognition group for all of their support and
encouragement over the years: Jennelle Yopchick, Erienne Weine, Maureen Gillespie, Amy DiBattista, Darrell Penta, Tracy Lennertz, Xu Zhao, and Amanda Dupuis.

Lastly, but most certainly not least, I could not have imagined making it through the last two years without the support of my boyfriend, Clayton Smith. While our relationship started at an unlikely time and place in my graduate career, his unwavering support and encouragement has always helped to remind me that I could do anything I set my mind to. I can never truly express how much it means to me to have someone willing to help me pursue my own dreams. I can only hope that I can support him in kind as I start a new chapter in life with him.
Abstract of Dissertation

As human beings, we navigate a rich complicated social world every day. Separating people into any number of social categories allows us to make quick intuitive predictions about the properties that people will have, even if we know very little about them. Psychological essentialism proposes that we believe groups have an underlying essence which confers group membership and gives rise to the properties associated with that group (Medin & Ortony, 1989). These beliefs influence predictions we make about the origin of category properties (innate potential), the extent to which members of the same category will share properties (inductive potential), and our willingness to accept changes from one category to another. While previous work has examined the extent of essentialist reasoning about natural kinds such as animals (e.g., Gelman, 2003), the present work aimed to determine whether adults have these intuitive beliefs about social groups.

One consequence of essentialist thinking, innate potential, is a belief that group members, by virtue of their underlying essence, will inevitably develop properties associated with their group regardless of external influence. Previous work from Eidson and Coley (2014) showed that adults reasoning under a time constraint exhibited increased innate potential compared to those under a time delay. Experiments 1 through 3 expanded on this work to determine if adults would exhibit similar increases in beliefs about innate potential under cognitive load (Experiment 1), within-subjects for reasoning about gender (Experiment 2) or for other social groups (i.e., race, gender, religion, and political affiliation – Experiment 3). To assess innate potential, participants read two switched-at-birth scenarios in which children were born to members of one social group and raised by another. Then, under cognitive load or a time constraint or time delay, participants made predictions about the child’s behaviors and physical characteristics later in life. Under
cognitive load in Experiment 1, beliefs about innate potential of gender did not significantly increase. However, participants gave essentialist responses faster than non-essentialist ones. In Experiments 2 and 3, for all social groups examined, beliefs about the innate potential of social behaviors significantly increased under time pressure. This suggests that although beliefs about the innate potential of social groups may be suppressed at baseline, adults may hold underlying beliefs about innate potential for race, gender, religion, and political affiliation which becomes more apparent when forced to make quick decisions.

A second consequence of essentialist thinking is that due to the deep, category-defining nature of essences, we may be unwilling to accept changes from one category to another. Experiment 4 was an exploratory effort to determine what properties people believe must be changed in order to move from one social group to another. For race and gender, participants read about an individual who changed stereotypical category behaviors, physical traits, or both. For religion and political affiliation, participants read about an individual who changed characteristic behaviors, nominal classification properties (e.g., properties necessary for objectively identifying as Christian), or both. Participants then rated which social group the individual was more representative of before their change and after their change. Results showed that changes to physical properties of race and gender resulted in a greater change in category membership than changes to stereotypical behaviors. Conversely, for religion and political affiliation, changes to characteristic behaviors resulted in a greater change in category membership than changes to nominal properties. Thus, for all groups we find that there are critical properties that when changed force greater changes in category membership. This suggests that we may have differential beliefs about the underlying nature of social groups.
Experiment 5 examined another consequence of essentialist thinking - inductive potential. Inductive potential is the extent to which we believe two individuals will share a particular property based on their shared category membership. Participants in Experiment 5 learned a novel behavior or novel physical property about a person who belonged to four different social groups (a racial group, a gender group, a religious group, and a political group). They then rated how likely other individuals who either did or did not belong to the same groups were to share these novel properties. Participants were willing extend these novel properties for each category, but participants were more likely to extend a novel physical property if two individuals shared the same race or gender and were more likely to extend a novel behavior if two individuals shared the same religion or political group. Thus, adults’ willingness to say that individuals will share similarities is dependent on both the particular type of social category and property being reasoned about.

Further, based on the results of Experiments 1 through 3 we propose that essentialist reasoning occurs within a dual-processing framework. Given that essentialist reasoning may be suppressed at baseline and re-emerge under time pressure, we posit that essentialism may best be thought of as a fast, automatic Type 1 bias which is otherwise suppressed by more explicit Type 2 reasoning processes. Throughout these experiments we also attempted to identify individual differences which may help explain differences in individuals’ essentialist reasoning and point to specific triggers for Type 2 inhibition of essentialist responding. Results suggest that the interaction between individual differences and essentialist reasoning is incredibly complex with differing relationships between such individual differences and essentialist responding depending on the task and category being reasoned about.
Taken together these findings suggest that we not only view social groups as informative for making predictions about different individuals, but also believe that social group membership may determine what an individual is like. These assumptions allow us to make quick intuitive predictions about individuals in order to navigate a complex social world – whether that is determining the types of properties that people may have and why have them or identifying when we feel a change in category status has occurred.
# TABLE OF CONTENTS

Acknowledgments ii
Abstract of Dissertation v
Table of Contents ix
List of Tables x
List of Figures xi
General Introduction 1
Experiment 1 55
Experiment 2 65
Experiment 3 87
Experiment 4 106
Experiment 5 132
General Discussion 151
References 175
Appendix A 184
Appendix B 189
Appendix C 192
Tables 194
Figures 204
List of Tables

Table 1. Property pairs used in the Switched-At-Birth task for Experiments 1 and 2. (pg. 194)
Table 2. Predicted directions of association between individual differences and speeded essentialist responding or degree of change in essentialist responding. (pg. 195)
Table 3. Beta ($\beta$) coefficients for individual difference predictors in Experiment 2. (pg. 196)
Table 4. Properties used for the Switched-at-Birth task in Experiment 3. (pg. 197)
Table 5. Beta ($\beta$) coefficients for individual difference predictors in Experiment 3. (pg. 198)
Table 6. Summary of properties used for each category and transformation condition in Experiment 4. (pg. 199)
Table 7. Beta ($\beta$) coefficients for individual difference predictors for *race and gender* in Experiment 4. (pg. 200)
Table 8. Beta ($\beta$) coefficients for individual difference predictors for *religion and political affiliation* in Experiment 4 (pg. 201)
Table 9. Mean likelihood ratings for each social category by property and match conditions in Experiment 5. (pg. 202)
Table 10. Beta ($\beta$) coefficients for individual difference predictors in Experiment 5. (pg. 203)
List of Figures

Figure 1. Mean essentialist responses by property type and loading condition for Experiment 1. (pg. 204)

Figure 2. Mean reaction time by property and response type for Experiment 1. (pg. 205)

Figure 3. Mean essentialist responses by property type and timing condition for Experiment 2. (pg. 206)

Figure 4. Counts of individual response patterns in each timing condition for Experiment 2. (pg. 207)

Figure 5. Proportion of essentialist responses for each category and property type for Experiment 3. (pg. 208)

Figure 6. Counts for individual response patterns in each timing condition for Experiment 3. (pg. 209)

Figure 7. Differences in category judgments by transformation condition for race and gender in Experiment 4. (pg. 210)

Figure 8. Differences in category judgments by transformation condition for religion and political affiliation in Experiment 4. (pg. 211)

Figure 9. Participant response patterns by transformation condition for the category of race for Experiment 4. (pg. 212)

Figure 10. Participant response patterns by transformation condition for the category of gender for Experiment 4. (pg. 213)

Figure 11. Participant response patterns by transformation condition for the category of religion for Experiment 4. (pg. 214)
Figure 12. Participant response patterns by transformation condition for the category of *political affiliation* for Experiment 4. (pg. 215)

Figure 13. Example of a dyadic trial with the base figure (black female Christian Democrat) on the left and the target figure (white male Christian Democrat) on the right for Experiment 5. (pg. 216)

Figure 14. Difference in match and mismatch likelihood ratings by property and category for Experiment 5. (pg. 217)
Overview

As we interact with the world, we obtain and store vast amounts of information about the things we encounter from objects to people. By encapsulating information about the world into categories, we can more easily access our knowledge about different groups. Based on the knowledge we obtain, we can also make informed inferences about things we do not know or have yet to encounter like a new animal or a new person. The current research specifically focused on the ways in which we use categories to understand and make inferences about groups of people.

There have been many proposals about how we organize and use category information. These have included the classical view (i.e., categories have necessary and sufficient features which result in either/or category membership), probabilistic views (i.e., category membership is graded based on shared featural similarity), and theory-based views (i.e., category membership gives rise to category features). These accounts are generally outlined in the following sections followed by a summary of past work and a proposed framework for the current work examining the ways in which we might think and reason about groups of people.

Classical View

Dating back to the philosopher Aristotle in ancient Greece, the classical (and aptly named) Aristotelian view of categories was the first approach taken by modern cognitive psychologists. For years, this particular assumption about the way in which we structure our concepts shaped our understanding of cognitive processes and development.

The classical view’s central claim about conceptual structure is that concepts are defined by a set of both necessary and sufficient features. That is, the particular features of a concept or category must be present in order for an item to be considered a category member, but the joint
presence of these features is also required. For instance, a member of the category “triangle” must be a closed figure, have three sides, and have angles totaling exactly 180 degrees. Each of these individual features is necessary for an object to be considered a triangle. Thus, a triangle must have three sides. It must have angles that equal 180 degrees. However, any single feature, while necessary, is not sufficient to justify category membership as a triangle. For instance, a three-sided figure with angles totaling 180 degrees that is not closed could not be considered a triangle. Although it contains two of the necessary features for triangle membership, the figure lacks the sufficient conjunction of features to be considered a triangle.

However, the classical view need not be limited to an enumerated list of features. Logical modifiers can help to create seemingly complex conceptual structures without needing to meet long, stringent criterion lists with perfectly uniform category members. In the triangle example above, the modifier “and” is used to denote that all of the necessary features must be present to sufficiently warrant category membership. Another way to think of this would be to consider each feature in the definition to be separated by a logical modifier. Thus, a triangle could be conceptualized as a closed object and having three sides and having angles totaling 180 degrees.

Similarly, a concept or category may use other logical modifiers such as “or” to create more complex combinations of items which can all be considered category members. As an example, the concept of “soda” might be something that is a liquid and is carbonated and is sold in a bottle or a can. Using the “or” modifier allows otherwise variable items to be considered as members of the category. That is, a carbonated liquid in a can will be considered a soda as will a carbonated liquid in a bottle. These modifiers allow both for some variability in actual item features while maintaining the necessary and sufficient definitions that are the hallmarks of the classical view.
In this sense, concepts may actually be thought of as being constituted by their definitions. That is, under the classical view, the concept of a triangle can be understood to simply be the definition of “a closed shape with three sides and angles totaling 180 degrees.” Once the logical conditions of the definition are met, any and all category members should be treated as equally good and representative members. Since the necessary and sufficient features define the concept, extraneous or unrelated features should be ignored upon gaining category membership. That is, because the definition is constitutive of the concept under the classical view once the definition is achieved, category membership is affirmed and assured regardless of those extraneous features. This approach also results in binary category membership. Either a given object has the necessary and sufficient features required for category membership (in which case it is indeed a category member) or it lacks such a combination of features (excluding it from category membership).

Unfortunately, while this definitional view of concepts handles simple categories like triangles or nominal kinds (i.e., “things you do on Sunday”) with relatively few and easily articulated features well, it is decidedly less effective at capturing the full scope of how individuals think about and determine category membership for more complex natural or social kinds. While classic experiments do show that participants can learn that certain concepts are explained or constituted by definitional rules consistent with the classical view, the nature and consequences of such definitions inevitably pose a fatal flaw for this approach. That is, the notion that concepts are driven by relatively stringent definitions of necessary and sufficient features sounds plausible, but experiments like those by Hampton (1979) show that participants are often quite poor at articulating what those features should be. That is, while they can often provide necessary features (e.g. mammals must have fur), the appropriate combination of
sufficient features is often much harder to produce in a way that captures all known category members.

For instance, we can make a strict list of features that are necessary and seemingly sufficient for classification as a mammal – has fur and gives live birth and is animate and breathes – but this list obviously excludes several known mammals. Specifically, marine mammals such as whales and dolphins do not have fur in the same sense of land mammals while some mammals like the platypus or echidna lay eggs instead of giving birth to live young. Thus, we must either remove these particular animals from the category of mammal or revise the necessary features to accommodate them. Doing so might only leave the necessary features “is animate” and “breathes.” However, this raises the problem that the necessary features overlap with many other possible types of animal categories. Thus the category of mammal must exclude known cases or be absorbed into a larger, less informative category of simply “animals.”

As illustrated above in the example of the whale and platypus, another major issue with the classical view is that it fails to take into account that individuals’ often accept seeming outliers into their categories. This has the effect that some members of a category are construed as more typical than others. That is, a dog or cat might be very typical of mammals while whales or platypus might be viewed as comparatively atypical. However, there is no accounting for this in the classical view as all category members should be considered full-fledged, equal members upon meeting the requisite criteria. These issues surrounding the enumeration and combination of necessary and sufficient features as well as typicality and category membership led to the formation of new categorization models that were probabilistic in nature as opposed to the rigid, absolute criteria imposed by the classical view.
Probabilistic Views

Probabilistic models of categorization attempt to capture typicality effects and graded category membership by moving from a necessary and sufficient lists of features to a system in which goodness of fit compared to an abstract representation helps to determine category status. In the most basic sense, any given entity can be compared to such a representation and be considered a better or worse category member depending on how congruent it is with the given representation. Highly congruent entities would be considered category members, highly incongruent entities would not be considered category members, with a broad range of “iffy” category members in between. While a cut-off point would ultimately determine the extent to which an entity could be unique and different from the category before no longer being considered a member, such a spectrum of congruency to the group representation can account for the typicality effects and fuzzier category boundaries which ultimately led to reexamination of the classical view. Though endless debate has ensued regarding the form of abstract representation that is most appropriate for determining category membership, for the purposes of this proposal I will briefly outline the two main probabilistic accounts – the prototype and exemplar models – to illustrate mechanisms of probabilistic categorization decisions while remaining ultimately agnostic between the two.

Prototype View. One form of abstract representation against which we may make categorization decisions is a summary representation of a category known as a prototype - an average across all category members (e.g. Posner & Keele, 1970; Rosch & Mervis, 1975). If enough features are shared between an entity and the abstract prototype, then category membership may be assumed (see Malt, 1989; Medin, Altom, & Murphy, 1984). Ultimately the
prototype need not actually resemble any given category members as it is an abstraction which averages across many known category members and features.

For example my prototype of bird may share many generic features with birds I have previously encountered (e.g. a duck, an eagle, and a robin), but may average across those or other features. For instance, taking into account the size of the birds I have seen, it is likely that my prototype is of a size closer to a duck (due to averaging across the larger eagle, the smaller robin, and average duck). Further, my prototype likely has a beak and wings just as all three exemplars do. Averaging across many other features, it is possible that the resulting prototype would closely resemble but not match any known bird. Thus, if I encounter a new entity like a sparrow and must decide if it is also a bird, I can compare the sparrow to my bird prototype. I would note that the sparrow, while not exactly like the prototype, shares many similar features including a beak and wings, while only slightly differing in size and foot shape. As such, I am likely to conclude that it is indeed a bird.

**Exemplar View.** In contrast to the notion of an averaged prototype, the exemplar view posits that individuals retain some number of categorical exemplars which are individual instances of category members encountered and accumulated in memory over time to which any new entity may be compared to determine category membership (Medin & Schaffer, 1978). In order to determine category membership, similarity calculations are made to determine the extent to which any new case is similar or dissimilar to existing categorical exemplars both of the target category and of non-target categories. That is, the exemplar view is not only concerned with how similar an entity is to a category’s existing exemplars but how dissimilar it is from the known exemplars of other categories.
As in the prototype view, this ultimately creates a spectrum of goodness-of-fit to the category for any given entity. Cases which are highly similar to and share many features with existing category exemplars would be readily accepted as good members of the category while cases which are highly dissimilar to category exemplars (and presumably similar to the exemplars of a different category) would be readily rejected as poor members of the category. Between these two ends of the spectrum, cases would differ in the degree to which they were found similar to existing category exemplars resulting in some cases being viewed as better representations of the category while others, though still perhaps deemed category members, would be seen as less typical or ideal category members. Some exemplar models (Medin & Schwanenflugel, 1981) further contend that high levels of similarity to a few category members will ultimately result in a more typical or good category member than moderate levels of similarity with many category members.

Returning to the bird example above, a person is likely to have many exemplars of birds retained in long-term memory collected over the course of many years including robins seen in the yard, ducks at the park, penguins at the zoo, or a pet cockatiel that your friend kept at home. If a new small, winged creature with a beak were encountered, these and many other exemplars would be quickly recalled for comparison against this new animal to determine if it should be deemed a bird as well. In comparison to a robin once viewed in your yard, this new creature may be similar in size and share features like the shape of its leg and beak while being slightly different in color. Despite the color difference, given the high rate of similarity among the other features, a high similarity score may be assigned to the robin-creature pair. A similar calculation would then be performed against another exemplar such as a duck recently seen at the park. Those calculations may show that while both animals have wings, feathers, and a beak, they
differ significantly in size, the shape of their feet, and the actual shape of the beak. This mix of shared features might result in a moderate similarity score. Similarity calculations would continue to be performed not only with other bird exemplars (showing the animal to be dissimilar to a penguin viewed at the zoo due to highly differing body shape, size, and color) but against exemplars of other categories (showing the animal to be highly dissimilar from all exemplars of cats, dogs, horses, etc.). Once these similarity calculations were complete, given the overall similarity to bird exemplars and dissimilarity to exemplars from other animal categories, we would come to the conclusion that our newly encountered animal is indeed a bird.

**Classical versus Probabilistic Categorization.** Although the prototype and exemplar views certainly differ in their mechanisms for determining the category membership of new entities (one relying on an abstract average or summary representation across category members and features, while the other makes direct comparisons to known individual category instances), both rely on similarity calculations between some representation of existing category features and a new individual. Similarity of features can be both very high and very low with a cutoff point for inclusion into the category being somewhere in-between. Any new individual with high similarity to existing category features would be viewed as a very good members of the category while those which are still somewhat similar, but much closer to the cut-off point, would still be included as category members but be deemed relatively poor or atypical of the category. It is this spectrum of typicality and graded membership that provides a major advantage to these probabilistic models of categorization over the classical view.

While the classical view relied on a set of defining and necessary features, the probabilistic models allow for the notion that some members are better examples than others (based on how similar or dissimilar they are from the given category representation). A wide
variety and instantiation of features is allowed by removing the importance of any single given feature and instead focusing on the overall similarity to many features. In this way, allowing for a range of typicality in category membership and avoiding the need for a defined set of necessary and sufficient features avoid the major issues with the classical view.

However, the probabilistic approaches are not without their own shortcomings. In particular, these approaches often fail to explain categorization decisions that occur when featural similarity and category membership are in conflict. For instance, both Keil (1989) and Rips (1989) demonstrated that changes in the appearance of an animal do not necessarily lead participants to recategorize it. That is, even though the features across which the probabilistic models would normally calculate similarity ratings for categorization may be similar to a different category following some transformation, an individual may often continue to classify it as part of its original category as opposed to the one that fits its new appearance. This suggests that calculations of featural similarity alone are not always enough to determine category membership.

**Knowledge-Based View**

Although the probabilistic views tackled many of the issues surrounding the classical view of categorization, they are not without their own shortcomings. The probabilistic accounts merely require enough known instances of a category in order for similarity calculations across features to be made to assess category membership for new cases. The knowledge-based approach, however, suggests that principles about how the world works can guide our categorization and inferences even when that knowledge is in conflict with featural similarity generally.
As Murphy and Medin (1985) argue, probabilistic accounts at times seem insufficient for explaining the range of categorization and inference phenomena that we can observe when dealing with rich, real-world stimuli. They propose that either intuitive or explicit theories about the ways in which the world works can constrain our categories and concepts by helping to form coherent links between features. This may be accomplished either through explicitly gained knowledge (i.e. domain-specific expertise) (e.g. Murphy & Wright, 1984; Proffitt, Coley, & Medin, 2000) or naïve theories about how the world and different domains function (e.g. Keil, 1989; Wellman & Gelman, 1992).

Some examples of the phenomena explained by such knowledge-based approaches include transformation of category members, physical differences between adults and offspring of a category, as well as differential reasoning patterns associated with expertise. For instance, Keil (1989) showed that children continued to assert that an animal such as a zebra should be classified as a zebra even if it was placed in a lion costume or physically transformed to appear like a lion. Based on probabilistic models, we would predict that after a transformation had occurred wherein the features of the zebra actually align much more closely with those of lion exemplars or prototypes, it should be reclassified as a lion with inferences about its behavior or appearance made accordingly. However, if people have an underlying assumption that there is a deeper nature to animals which cannot be affected by such a physical transformation, this might prevent recategorization even in the face of perceptual dissimilarity.

The same might be true in the case of parents and offspring who lack relatively few physical markers (Gelman & Wellman, 1991; Mervis, 1987). Even though parents and offspring may share relatively few physical similarities, guiding biological principles about the ways in which offspring generally stick close to their parents, tend to eventually resemble them, and may
inherit their qualities can lead children to include the two animals in the same category despite a lack of featural similarity. Based on the probabilistic models, because a young animal may share relatively few features with its adult parent (and perhaps that an offspring may in fact more strongly resemble a different type of animal), the two entities could be placed into different categories. Yet, if underlying beliefs about the nature of the world, of animals, or biology are taken into account, a different set of decisions may be expected, especially in the face of featural dissimilarity or uncertainty.

These beliefs about the underlying nature of category membership may seem disparate, but many of these notions can be united under the umbrella of psychological essentialism – a belief in an underlying essence which gives rise to the features of a category and defines category membership (Medin & Ortony, 1989). The following section and the remainder of this proposal will explicate and consider further evidence for psychological essentialism and questions regarding its place as a unifying knowledge-based approach to categorization and inductive inference.

**Psychological Essentialism**

**General Theory.** While essentialism has its roots in the philosophical traditions of Plato, it has been resurrected in several forms making its way into fields such as anthropology and psychology. In its original form, essentialism was a metaphysical claim about the nature of objects, individuals, and the world. First introduced in Plato’s *The Republic*, the famous allegory of the cave tells a story of people chained in a cave for their entirety of their lives with a fixed view of a single wall unable to view anything else in the cave or outside world. Every day, people and animals behind them passed between a fire and the wall casting shadows. From this perspective the cave dwellers consider the shadows to be constitutive of the nature of other
people and animals (Plato, 1992/380 B.C.). That is, the shadows would be perceived as the people and animals themselves, though objectively the shadows were merely the result of actual people and animals standing between a fire and the wall. Thus, the actual essence or true nature of other individuals and animals are unknown to the cave dwellers who have only seen an ambiguous representation of people and animals through shadows, though they would be unaware of this fact until they were released from the cave. On Plato’s view then, every object, person, or animal has a true underlying nature or essence which can be discovered and understood, informing us as to its true characteristics.

While Plato’s essentialism makes metaphysical claims about the true, objective nature of people and objects, psychological essentialism instead suggests that the human mind has an underlying belief in or assumption of the existence of essences, whether any such essence exists or not. Medin and Ortony (1989) are some of the first to lay out the idea of psychological essentialism as one devoid of metaphysics. While they choose not to debate the metaphysics of essences themselves, the authors are most concerned with people’s beliefs about the existence of essences. In their opinion, people tend to behave in a manner consistent with a belief in an underlying essence.

Medin and Ortony’s (1989) theoretical proposal is that beliefs about essences include the following: an essence is the causal agent which gives rise to various properties associated with a category, is present in all category members such that all members exhibit similar if not identical properties, and is non-obvious (or not easily identifiable via observable features) to a lay person. Thus the presence of an underlying essence can be said to confer category membership (i.e. if an animal has a bird essence, it will be considered a bird both having the underlying essence of bird, but also bird-like properties caused by that essence). Further, essences are considered non-
obvious and causally responsible for giving rise to more observable category-based properties. This allows for surface-level changes to be made to a category member without changing category membership (as in Keil, 1989 outlined above). That is, simply changing the appearance of an individual cannot change category membership as the individual would still contain the deep underlying essence (making it a category member) even though the physical features typically associated with that category may no longer be present. Thought of another way, while transformations may effect changes in appearance or other surface-level properties, presumably several changes could be made without actually affecting the underlying essence. As such, category membership can be considered to remain constant even in the face of shallow transformations as long as the essence remains in tact.

Gelman (2003, 2004) further suggests that psychological essentialism may best be thought of as an early emerging cognitive bias (as young as two years old in Gelman & Coley’s, 1990 work). That is, even with little explicit knowledge about the world, young children appear to make categorization decisions which are consistent with essentialist beliefs. Without explicit knowledge, Gelman (2004) proposes that essences are most probably inferred by the presence of a category label. That is, upon hearing that two particular animals belong to the group “bird,” a person might infer that the label “bird” signals the presence of an essence, and since the two animals belong to the group “bird” they likely share the “bird essence.” As Medin and Ortony (1989) point out, this is not to suggest that someone might not be able to identify an essence should one exist. Indeed, they suggest that someone with appropriate expertise about a given category or domain might be able to identify the true (or perceived to be true) essence of a given group. Yet, lay people may use the category label as a placeholder to indicate the presence of an
essence without necessarily understanding what properties that essence causes or being able to identify the essence itself.

Gelman (2003) is careful to note that essentialist beliefs should not necessarily be considered explicit beliefs. Given the age at which these phenomena arise in children, it seems unlikely to suggest that explicit essentialist theories are held and guide inductive inference. However, older children and adults can and do explain their inference choices often through formal (e.g., “She plays with tea sets because she’s a girl.”) or associational explanation (e.g., “She plays with dolls because girls like dolls more.”) (e.g. Taylor et al, 2009; Samarapungavan & Wiers, 1997). However, given the early emergence of essentialist inferences, such explanations may simply be post-hoc justifications for responses rather than explicitly-held principles of decision making.

In summary, based on both the knowledge-based effects outlined above as well as Medin and Ortony’s theoretical outline of psychological essentialism essences are assumed to be deep, causal agents which give rise to the identifying properties of a category when the essence is present. These essences can be found in all members of a category, thus leading to a belief in similarity or near-uniformity of group members. These essences may be transferred from parent to offspring through reproduction, allowing features tied to the essence to arise through natural development, if not immediately. Because of their deep, underlying nature, surface-level changes to the physical appearance of an entity should not necessitate a change in category membership even when featural similarity favors inclusion in a new and different category. Similarly, even in the face of perceptual dissimilarity, category membership can assumed to be shared if the two disparate entities are assumed to share the same essence. Further, the use of category labels as
points of inference can signal essentialist reasoning as laypeople may use category labels as an essence placeholder when the exact nature of a category essence is unknown.

It should be noted that if we are indeed influenced by an essentialist bias in the ways outlined above, the resulting category structures and inferences turn out looking very much like a belief-driven classical view. Under psychological essentialism, essences demarcate category membership such that those who are believed to contain the category essence are considered category members and those who lack it are not. Further, since category features are tied to and caused by the essence, all members who contain the essence should exhibit these properties. This can look very much as though there is a coherent set of necessary features for category membership but which are actually tied to the category essence. And since those who lack the essence should similarly lack those properties, it could appear that when the necessary features are absent (as is the essence) then category membership is rejected. The main difference here is that the essentialist view encompasses specific assumptions about the ways in which an essence can function, be transferred, and is related to category properties while the classical view merely requires that a logical set of requirements be met for inclusion or exclusion in category membership.

In the following sections, empirical evidence for the essentialist account will be discussed. This will help to show that essentialist assumptions about categories guide our categorization and inference decisions both throughout development but into adulthood as well. Further, these essentialist assumptions can be seen in both natural and social kinds – two meaningful distinctions we can make between the greater natural world and the ways in which we can classify the human world. To date, the vast majority of our understanding of essentialist reasoning has been derived from work tracking the developmental course of essentialism in
children. In many cases, adult participants were simply not included in these studies. Thus, although the current work focused on the extent of adults’ essentialism, the work reviewed below reviews the general framework of essentialist assumptions through which we examined essentialist reasoning in adults.

**Essentialism in Natural Kinds**

Given the theory and implications of essentialism above, essentialism can explain several seemingly disparate phenomena found throughout the categorization and inference literature. Specifically, a belief in essences can help to explain these phenomena in a way that probabilistic accounts cannot.

**Category-based inference despite featural dissimilarity.** To the extent that category members can be assumed to share the same underlying essence which gives rise to category-based properties, then we can make assumptions about the properties that new individuals who are deemed to also be category members will have (e.g., Gelman & Coley, 1990; Gelman & Markman, 1986). If this assumption were driven only by shared surface-level features, then we should see little or no extension of properties to individuals who do not share enough featural similarity with existing category members. However, if the assumption was driven not by external features, but by deep, non-obvious links between category members, then simply knowing an individual is part of a category, regardless of how similar it is to other category member, would make category-based inferences possible.

Categorical inferences may best be defined as the use of previously learned information about category members to help understand or predict the behaviors or properties of new instances. That is, based on my existing knowledge of cats (i.e. cats are generally small, furry, four-legged, meowing mammals), I am likely to classify any new furry, four-legged creature that
meows as also being a cat. Further, if I saw an animal which I knew to be a cat, I would likely assume that it could meow even if I had not heard it do so. Thus, categorical inferences allow us to generalize from what we already know and apply it to new, uncertain cases.

In the example above, the classification of a novel animal as a cat is at least aided by perceptual similarity or perceptual features. That is, features which are known to be associated with cats (e.g. furry and four-legged) are visibly accessible. Yet, when making such inferences, we may encounter new instances which are perceptually similar to the exemplars or prototype of one category even though category membership should not actually be shared. Take for instance, the stimuli used by Gelman and Markman (1986). Children were presented with a triad of animals matching either in perceptual or conceptual similarity. For example, children might see a bat, a seagull, and a flamingo. The bat and seagull are perceptually similar with both having long wing spans and similar body structure. However, the seagull and flamingo, while sharing very few perceptual features, both belong to the category of “bird.” If children simply relied on perceptual similarity or shared features between exemplars or an abstract prototype, we might expect that inferences would be made from the bat to the seagull. However, Gelman and Markman’s data suggest that both children and adults rely not on perceptual similarity, but category membership or conceptual similarity to make inferences about novel properties when category labels are present. That is based on premises such as, “This [flamingo’s] heart has a right aortic arch only,” and “This bat’s heart has a left aortic arch only,” both children and adults would be more likely infer that the seagull would have a right aortic arch because it belongs to the same category – “bird” - as a flamingo even though it shares perceptually similar features with the bat.
These findings are similar to those of Carey (1985) who found that children used their biological knowledge of human non-obvious properties to enrich their inferences about other animals. That is, children often ignored external similarity in favor of making inferences based on the extent to which they viewed the animals as similar to humans. If surface similarity alone were the driving component of children’s inferences, knowledge of human non-obvious properties (such as humans have spleens) would not be applied to other species (i.e. to make the inference that dogs have spleens).

Relatedly, Gelman and Markman (1987) showed that children’s inferences differentially changed based on the label ascribed to a perceptually ambiguous instance. That is, when a figure that had features congruent with both a leaf and an insect was presented, children ascribed different properties to it based on the label presented. If they were told it was a leaf, they would ascribe leaf-like properties to it (though it still had features similar to an insect), but if they were told it was an insect, they would ascribe insect-like properties to it (thought it still had features similar to a leaf). Further, children may make category-based inferences about novel objects or animals when labeled with either familiar or novel category names (Davidson & Gelman, 1990), when category membership and perceptual appearance are in conflict (Gelman, Collman, & Maccoby, 1986), and when little or no perceptual information is available (e.g. Diesendruck & haLevi, 2006). Further, while the majority of these studies have focused on four- to six-year olds, similar findings have been noted in two-year olds, though only in the presence of category labels (Gelman & Coley, 1990; Jaswal & Markman, 2002).

Taken together, these findings suggest that children can and do ignore perceptual similarity when making inferences. Instead, category membership or conceptual similarity is used to extend underlying, non-obvious novel properties from one instance to another, even
when perceptual features are ambiguous or in direct conflict with category membership. Such inferences and categorization decisions are not easily captured by exemplar and prototype theories. Under those theories, perceptual ambiguity or perceptual conflict should cause inferences to fail or to shift in favor of perceptually similar instances as opposed to conceptually similar ones. That is, an evaluation of salient features should determine that high matching of perceptual features should constitute shared category membership and thus promote inferences. Yet, we see that conceptual similarity often overrides perceptual similarity, suggesting that some other guiding principle must be affecting reasoning.

**Constancy or Resistance to Change.** If deep, underlying, non-obvious properties and essences played no role in our categorization decisions or inferences, then substantial changes to surface-level features should also effect changes in category membership and willingness to make category-based inferences. However, while the outward appearance of a person or animal might change, this is not necessarily reflected in the way in which we reason about that thing or ascribe category membership. While the studies outlined above pit perceptual similarity against conceptual similarity across instances, evidence has also been accumulated which suggests that changes in perceptual similarity within an individual do not necessarily cause shifts in category membership or category-based inferences.

One of the best examples of this comes from Keil’s (1989) work on children’s understandings of changes in appearance as they relate to property expression. For instance, when children were told that an animal has been placed in a costume (i.e. a lion in a tiger costume), they still identify it by its original category membership. That is, simply changing outward appearance does not necessarily affect a change in understanding of category membership.
Similarly, Rips (1989) showed that transformations in appearance based on developmental maturation (i.e. appearance changed as an animal grew) did not lead adult participants to recategorize the animal. Rips presented a scenario to participants wherein an animal began its life looking similar to an insect but changed in appearance to eventually look more like a bird. When this was identified as a natural outcome of development, participants reported that the category membership at the end of development was the same as it had been at the beginning of development, even though the two stages differed greatly in similarity. Similar expectations exist for humans (Rosengren et al, 1991) such that children are willing to make inferences between babies and adults consistent with the notion that changes in appearance do not necessitate changes in category membership.

Similar findings reported by Hampton, Estes, and Simmons (2007) suggest that at least some adults may also preserve category membership even in the face of accidental transformations or mutations. Using Rips’ paradigm, Hampton et al told participants that either a chemical causes a change in appearance or those changes were due to development. Under both scenarios, many participants maintained category membership regardless of how or why the change occurred indicating that regardless of circumstances or type of change, category membership was viewed as stable and unchangeable even by external agents and when an individual had taken on a completely different set of perceptual features (though this conflicts with Rips’ findings that “accidental” changes lead to changes in category membership).

However, “deeper” changes to individuals can and do necessitate changes in property expression and category membership. For example, Keil (1989) told children that scientists changed both the outer appearance of an animal to that of another species as well as replacing its insides to match those of the other species. Once the insides matched, children inferred that
category membership had changed. Similar findings reported by Gelman and Wellman (1991) suggest that even simply removing insides without necessarily replacing them (i.e. removing the blood or bones from a dog) will change membership, though simply removing outside without replacing them (i.e. remove the dog’s hair) will not change category membership. Such changes do not simply affect children’s reasoning, however. The changing of insides in Keil and Gelman and Wellman’s studies is analogous to the real world medical procedure of having an organ transplant. A study on adult patients found that 34% of heart transplant recipients felt that they may have taken on attributes and properties of their donors, including changes in appetite and acquisition of new hobbies (Inspector, Kutz, & David, 2004). That is, in receiving a donor organ (new insides) many heart transplant patients saw themselves as aligning with the previous donor, changing their outward properties such as appetite and hobbies to conform to those of the donor. Thus, while simply making external changes is not sufficient to force a change in beliefs about category membership, making deeper changes to non-obvious properties of an individual, especially when those changes match a different category can and do effect such changes.

Taken together these findings suggest that both children and adults do not rely solely on outward appearance or the present state when determining category membership and making inferences. However, the findings do suggest that deeper changes to category members (i.e. changing “insides”) do affect judgments about category membership, shifting instead to toward the category and properties of the new insides as opposed to the old. Just as in the case of perceptual ambiguity outlined in the previous section, exemplar and prototype views fail to account for these types of categorization decisions. Under those views, once an individual’s appearance has shifted to match the typical appearance of another category, when calculations are made over the salient features, category membership should also shift to match the
individual’s new appearance. Yet, we see that participants consistently reject the notion that surface level changes necessitate a change in category membership. Instead, it is the consideration of changes to deep, non-obvious properties that can effect a change in category membership suggesting that principles related to non-obvious properties override reliance on surface-level featural similarity in reasoning.

**Inheritance and Innate Potential.** Aside from category structure and category features themselves, naïve understandings of the biological nature and principles of categories and their members can also influence our categorization and inferences. For instance, the presumed link between parents and offspring and beliefs about biological inheritance also inform inductive inference even in the face of other factors such as perceptual similarity or environmental influence. That is, even when two members of a category are perceptually dissimilar, the understanding of a conceptual link via kinship can allow inferences to be made. Beliefs in such biologically-linked, non-obvious traits are consistent with the notion of an underlying essence giving rise to various category-based properties. Further, these beliefs about one category member signals its likely existence in all category members. Such a relationship may not be limited to members who are biologically related – although a biological trait may be passed from parent to offspring - but instead it is diffuse among category members, presumably even when there are perceptual differences between those members. Thus, the presence of non-obvious properties in a single individual can be a signal that other category members should have similar if not identical non-obvious properties, even if the category members are perceptually dissimilar (either in the case of parents and offspring or variable group members)

As an illustration, Springer (1992) presented children with triads of animals with a target being related to the two premises either by kinship or via perceptual similarity. That is, animals
may be known to be biologically related but perceptually dissimilar or biologically unrelated but perceptually similar. However, even in the face of perceptual similarity, children made inferences about novel properties based on kinship, exhibiting an understanding that properties may be inherited via biological reproduction. Solomon et al (1996) further found that by age seven, children mostly make inferences about shared biological properties (but not shared beliefs) along kinship lines. This suggests that other, non-obvious relationships between category members such as kinship can override reliance on perceptual similarity when making inferences suggesting a deeper root cause of making inferences consistent with beliefs in an essence.

However, these inferences are not limited to kinship relations. Springer and Keil (1989) found that children believed properties which were considered inborn and present in one category member such as an abnormality were relatively likely to be exhibited by nonkinship members of the same species. This is opposed to properties which may be acquired during the course of one’s lifetime and are presumably not passed on via inheritance even to children. This speaks to the innate potential of categories. That is, if a property can be ascribed to inheritance, even for some members, the potential for property expression exists across all category members. This suggests that a deep relationship like a shared essence exists between category members.

The innate potential of categories has often been demonstrated by using a switched-at-birth task like the one developed by Gelman and Wellman (1991). The switched-at-birth task proposes a scenario in which an individual is sent to live with members of a different category shortly after birth. Participants must then decide whether or not later in life the individual exhibits properties based on birth category or consistent with the birth parents (signaling a belief that the individual’s category essence determines its traits) or based on adoptive category or
consistent with the adoptive parents (signaling a belief that the category essence is either absent or not determinative of traits). Gelman and Wellman’s findings suggest that children maintain a preference for the birth category when making inferences (i.e. an individual will exhibit properties consistent with the birth category, not the adoptive category). Thus, property expression in an individual is not simply changed because of the environment. Instead, the underlying link between the individual and the category (and its members) persists leading to category-based (and not environmentally-based) inferences about the individual’s properties.

Similar findings by Gelman and her colleagues and other have suggested that this innate potential extends to behavioral and physical properties of categories which might be assumed to be biologically inherited (e.g. Taylor et al, 2009; Johnson & Solomon, 1997) as well as to non-biologically inherited properties like language (Hirschfeld & Gelman, 1997). Further, innate potential seems to exist both for reasoning about animals (e.g. Gelman & Wellman, 1991; Taylor et al, 2009; Johnson & Solomon, 1997) as well as for racial categories (e.g. Hirschfeld, 1996; Rhodes & Gelman, 2009) and gender categories (e.g. Taylor, 1996; Taylor et al, 2009). It should be noted, however, that adults make attenuated inferences of innate potential such that category-based inferences are made for the physical and behavioral aspects of animals (e.g. Gelman & Wellman, 1991; Rhodes & Gelman, 2009; Taylor et al, 2009) but only for physical properties of people (e.g. Rhodes & Gelman, 2009; Taylor et al, 2009). Given that children ascribe innate potential to the both physical and behavioral properties of both animals and people, this suggests that adults’ attenuation of innate potential when reasoning about people is a function of age rather than a domain-specific phenomenon. This attenuated pattern of inference about people and its importance will be further explored in subsequent sections.

Summary
Taken together these findings suggest that not only are biological relations believed to outweigh perceptual similarity, but this belief holds even under inhospitable circumstances. That is, category members are believed to exhibit category properties even when immersed in an environment full of other category members exhibiting different properties. The failure of perceptual similarity in the face of such relations again questions the extent to which the prototype and exemplar views can account for these findings. The deep, non-obvious link of kinship is at best a single feature to be considered when trying to calculate category membership. Yet, biological relationships can override perceptual similarity in categorization decisions as well as overriding other plausible considerations like the effect of the environment. Thus, such deep, non-obvious properties as well as the principle of biological relatedness are consistent with beliefs in a deep, underlying essence which both signals category membership but also is the cause of category-based properties. Further, this essence can influence the reasoning system even in the face of equally useful information like perceptual similarity and environmental influence.

**Essentialism in Social Kinds**

While it seems clear that both children and adults (to some degree) exhibit essentialist patterns of reasoning which go above and beyond typical prototype and exemplar models of categorization, different domains (i.e. between types of categories – natural kinds, social kinds, and artifacts) have been shown to have differing degrees of essentialism. Although differences in essentialism have been noted certainly between natural kinds and artifacts (e.g. Rhodes & Gelman, 2009) work by Bloom (1996) has suggested that this may be due specifically to understanding the intent of use for artifacts. However, in tasks which demonstrate essentialism such as the SAB, differences have also been noted in the extent to which natural kinds and social kinds are used for category-based inferences (e.g Taylor et al, 2009, Rhodes & Gelman, 2009).
Although social kinds are specifically groupings of people, human beings are of course still natural living kinds (though various social kinds may not be ontologically natural). This begs the question to what extent social kinds might be treated as natural kinds or essentialized to the same extent as natural kinds, and if they are not, what factors might help to explain this discrepancy. Further, it should be examined whether or not such a discrepancy is uniform for all social categories or simply ones which are deemed less natural than others (i.e. race or gender versus religion or occupation).

Social kinds are particularly interesting in that while natural kinds are often philosophically viewed as “carving nature at the joints” (e.g. Kornblith, 1993 in the tradition of Plato) or representing true differences between things in the world, social kinds might be thought of as differing in the degree to which they are naturally occurring. That is, many classifications of people represent rather arbitrary and myopic ways of carving up the social world. This is not to say that the categories themselves are meaningless. For instance, there are indeed occupations like “doctor,” “teacher,” and “flight attendant” which account for a certain number of individuals engaged in similar activities. There are people who go to medical school, get medical degrees, and practice medicine in hospitals and clinics. Thus the occupation “doctor” is not without meaning. However, “doctor” as a social category is not the same as say the category of “zebra” or even “human” in that its individual members are exceptionally unlikely to share inherent, natural properties because of their membership in their occupational category. That is, physical appearance, behaviors, or preferences are often not attributable to the social category of occupation, but rather may be more causally related to individual’s status as human beings.

Further, most individuals belong to multiple social categories. A person can be both male and Catholic, a doctor and heterosexual, African-American and a music lover. The infinite
permutations of social category membership, coupled with the lack of causal power by social
categories to impart intrinsic properties to their members suggests that social kinds do not
represent a true carving of the social worlds in the way that philosophers would argue that
natural kinds do (e.g. Kripke, 1980). However, they may still be useful for making predictions
about others’ behaviors and appearance and thus are not necessarily meaningless, though they
may be qualitatively different from natural kinds.

Evidence for Essentialism in Social Kinds

While comparatively less research has been conducted on essentialism in social
categories, much of the research that has been conducted has focused on three main social
categories: race and ethnicity, gender, and religion. While in no way exhaustive of all possible
social categories, these initial investigations of social essentialism provide early evidence that the
social domain is a potentially rich area of study to both better understand which social categories
we might essentialize as well as the mechanism by which we might essentialize those categories.

Race and Ethnicity

Although race certainly has biological components, including a genetic basis for skin
color, once we begin to ascribe behaviors and preferences to the category of race, it moves out of
the realm of a biological or natural kind category to the realm of a social category. That is, once
we begin to use race or ethnicity as a means of predicting the behaviors of other people, it can no
longer be simply thought of as a biological category. Further, with skin tone (as a marker for
race) being a relatively salient characteristic, it has been one of the starting points for
investigators trying to understand whether and how people use social categories to make
inferences about others.
Much of the work on essentialism in race has been motivated by Hirschfeld’s (1995, 1996) work on children’s essentialist beliefs about racial categories. For instance, Hirschfeld (1996) demonstrated that children viewed racial groups as constant across the lifespan. Children were presented with triads depicting an adult of a particular skin color (e.g. black) as well as various other possible categories (e.g. occupation, body type), and two children matching either by race or along one of the other category dimensions. Participants were then asked which of the two children represented the adult when he or she was younger. On the vast majority of trials, children chose the racial category match over other category matches. That is, children predicted racial constancy such that individuals should belong to the same racial category during adulthood as they did during childhood. Further, this distinction was more important to children than trying to keep occupation or body type constant across time. This suggests that children viewed race as a more stable category over time and that other changes in appearance or occupation did not necessitate changes in racial category membership. Thus, at least compared to the other categories used in Hirschfeld’s experiment, race was relatively more essentialized such that participants expected race to persist over time allowing for inferences to be made between an adult and a child member of a particular category, in spite of their perceptual dissimilarity.

Similarly, Hirschfeld (1995) presented children with a switched-at-birth task wherein babies of two different families – one white and one black – were incorrectly sent home with the wrong family from the hospital. Participants were then presented with pictures of adolescent children who either did or did not match the birth parents in racial category. Participants were asked to determine which child belonged to the parents. If participants only believed that who had raised the child mattered, then the racial non-match would have been the more appropriate answer. However, if racial categories were determined at birth and the racial essence had been
imparted to the baby influencing its category membership while being unaffected by environmental influence, the racial match would be the appropriate answer. Indeed, children as young as four overwhelmingly chose the racial match, suggesting that racial category membership is viewed as not simply due to environmental or social influence but rather it is viewed as something inherent from birth. This suggests that children reason as though there is a racial essence which both confers category membership and affects category property expression.

Similar findings by Astuti, Solomon, and Carey (2004) showed that children of the Vezo tribe in Madagascar overwhelmingly supported the innate potential of ethnic tribe membership in the same way that children appear to think about race. That is, when presented with a switched-at-birth task, Vezo children usually reasoned that a Vezo child who had grown up with parents from a neighboring tribe would maintain their Vezo identity and category membership in spite of the environment in which they had grown up. This was in contrast to Vezo adults who overwhelmingly supported a socialization method of categorization. That is, to Vezo adults, category membership of Vezo tribe members was determined by presence in the tribe and engaging in Vezo activities, not by birth. Thus, when presented with a switched-at-birth task, Vezo adults reasoned that a Vezo-born child raised by members of a neighboring tribe would no longer be Vezo and, similarly, a child from a neighboring tribe raised by Vezo parents would be considered Vezo. This suggests that Vezo children have an essentialist understanding of ethnic tribe membership such that tribe membership at birth is the most important for determining category membership as opposed to the current tribal setting. Yet, as Vezo children age, this essentialist understanding appears to diminish and in fact reverse in adulthood.
Other work by Diesendruck and his colleagues (e.g. Birnbaum et al, 2010; Deeb et al, 2011; Diesendruck & Haber, 2009; Diesendruck & haLevi, 2006) has shown that children in Israel view ethnic groups like “Jews” and “Arabs” as meaningful for making inferences. For instance, Diesendruck and haLevi (2006) showed young children triads of members of different social groups including ethnicity who also had a particular psychological characteristic (e.g. “shy”). That is, children might have seen two individuals such as a shy Arab and an outgoing Jew. Participants then had to project a novel property (e.g., “likes to play zigo” or “wants to be a mashitz”) to a target figure (e.g. a shy Jew) based on either social category status or psychological category status. Thus, children could choose to extend the novel property based either on a psychological characteristic (i.e. from a shy Arab to a shy Jew) or based on ethnic category (i.e. from a outgoing Jew to a shy Jew). Children largely made property projections based on social category status as opposed to psychological trait category status. This implies that children in Israel find social categories like “Jew” and “Arab” to be more informative for making inferences than psychologically defining categories like “shy people.” Thus social category labels like “Jew” and “Arab” may be markers for an essence as outlined by Gelman (2004) allowing generalizations to be made between two unknown individuals who share category membership. However, it should be noted that adults largely made property projections based on psychological traits and not social category status, suggesting that groups like “shy people” may be more informative to Israeli adults for making inferences than ethnic groups.

It should be noted, however, that not all children in Israel follow this pattern. Indeed, young, religious, Jewish children tended to use ethnicity as a greater source of inference than their non-religious counterparts or Arab children (Birnbaum et al, 2010). This suggests that when reasoning about religious or ethnic groups, a participant’s religiosity and personal ethnic group
affiliation may influence otherwise essentialist reasoning. Further, Jewish children who are exposed to members of minority groups (i.e. Arabs) are more likely to mention and use ethnicity as a point of inference than Jewish children who are not exposed to other groups or Arab children (Deeb et al, 2011). However, Arab children who do not regularly interact with Jewish children at school are more likely to endorse the idea that ethnic categories are stable, unchanging, and heritable than their counterparts who attend schools with other Jewish children or Jewish children in general (Deeb et al, 2011). Thus, a participant’s personal experience and interaction with members of a racial or ethnic category may also influence their reasoning, similar to how development appears to attenuate essentialist responding.

Such results are similar to those found by Rhodes and Gelman (2009) which point to possible socialization or cultural differences in the expression of essentialist responding. Participants in their study were asked to judge whether or not a puppet’s beliefs about members of various categories were correct. Of critical interest were participants’ responses to presumably inconsistent category beliefs. For instance, Feppy, the puppet, may have said that a black boy and a white boy were members of the same category. Participants could either say that the puppet’s choice was correct, indicating that participants did not make distinctions between black and white people (because they were willing to group them together), or participants could say he was incorrect suggesting that category boundaries were being crossed. Although young children tended to accept the puppet’s grouping of members of different racial groups (suggesting that they did not find race to be a meaningful category for making distinctions between people), adolescents rejected this grouping (suggesting that they saw whites and blacks as members of different groups). However, when split by rural and urban environments, rural children tended to reject the puppet’s miscategorizations more than their urban counterparts and
such rejections increased with age. These results suggest that though a basic bias for use of race and ethnicity as a meaningful category and belief in rigid boundaries between racial groups may be present signaling belief in a racial essence, such a bias may be attenuated by other factors (e.g. age, majority / minority status, culture).

Overall, these data suggest that in general young children view race and ethnicity as an important category for making inferences. Further, racial category membership is believed to be constant over time and regardless of the current environment. Such phenomena are consistent with a belief in an essence which is present in all members of a racial group and causes similar properties to occur. Although there are some hints that older children and young adults hold similar beliefs, there are also instances in which adults fail to exhibit essentialist reasoning about race and ethnicity. For instance Vezo tribe adults showed a preference for using environmental context for inferring category membership as opposed to tribal membership at birth. Also, cultural context appears to play a part in the degree of essentialist responding as evidenced by the differences in the use of racial groups by urban and rural children in Rhodes and Gelman’s (2009) work. Thus, while race may be essentialized early on, the fate of racial essentialism across development is less clear. While it is possible that the essentialist bias for race actually is not present in urban children or adults, it is also possible that it is simply not being overwritten. This will be discussed in greater detail in following sections.

Gender

As with race and ethnicity, it is important to distinguish gender (a social category) from sex (a biological or natural kind category). Sex is a biological concept and is often, though not always, physically observable. A person is generally either born genetically male or female as reflected in their anatomy – though Blackless et al (2000) estimate that intersex conditions may
occur in up to approximately 2% of the population. Further, these genetic differences influence the amount of hormones such as testosterone or estrogen that are present in one’s biological system. Thus, sex encompasses biological properties which may affect or influence behavior, but the concept of sex itself does not necessarily include such behaviors.

On the other hand, gender is used to account for traits that sex cannot encompass. For instance, Bem (1981) proposed that gender could best be conceptualized as a schema – a set of predisposed feelings or traits – that a society has for each sex. Thus, the biological concept of the sexes translate into gender groups based on society’s schematic view of the sexes. This attachment of behaviors and psychological traits to the sexes thus results in a social category of gender which encompasses both the biological properties associated with sex as well as the social-behavioral properties stereotypically associated with males and females just as racial categories include both physical and behavioral attributes.

Children and adults seem to have a similar essentialist understanding of the social category of gender much like they do with race and ethnicity. For instance, gendered triad tasks presented to children indicated that participants were willing to extend non-obvious or novel properties based on gender category membership even when category membership was in conflict with appearance (i.e. a figure was labeled as a boy, though was female in appearance) (Gelman, Collman, & Maccoby, 1986). This suggests a belief in an underlying essence which persists even in the face of perceptual dissimilarity, allowing inferences to be made when a category label hints at the deeper, underlying shared essence.

Using a switched-at-birth task, Taylor (1996) showed that young children tended to make categorical inferences about males and females who had grown up in opposite-sex environments. That is, if a girl had grown up on an island with all men, young children were still likely to say
that she would exhibit stereotypical “girl properties.” This indicates that young children hold beliefs about the innate potential of gender category membership to influence property expression even in a dichotomously opposing environment vis a vis the continued presence of a gendered essence. However, as children entered adolescence and young adulthood, category-based inferences diminished for behavioral properties (e.g. “plays with tea sets” or “wants to be a firefighter”) though not for physical properties (e.g. “has a boy body” or “will grow up to have breasts”).

Work by Taylor, Rhodes, and Gelman (2009) extended these findings by comparing children’s essentialist responding about gender categories to their responses about natural kind categories (i.e. animals) on a switched-at-birth task. Their findings replicated those of Taylor (1996) such that younger children tended to make essentialist inferences about both physical and behavioral properties of gender and did so to the same extent for gender categories and animal categories. However, as children increased in age, essentialist responding about gender decreased overall compared to essentialist responding about animals, but as in Taylor’s findings, this effect was driven by a marked decreased in the amount of essentialist responding about gendered behaviors compared to essentialist responding about gendered physical properties (but see Eidson & Coley, 2014).

Rhodes and Gelman (2009) show similar results to those about race while mirroring the developmental pattern seen by Taylor (1996) and Taylor et al (2009). Overall, children rejected attempts to place males and females in the same group suggesting that children view “male” and “female” as two distinct and non-overlapping categories and each group has its own essence which defines its members and their properties. Such rejections generally decrease with age, however, suggesting an easing of the boundaries between gender categories (and decrease
reliance on a category essence) into adulthood consistent with the decrease in essentialist responding previously observed. However, as with racial categories, children of different age groups from rural environments exhibited high levels of rejection for miscategorizations of gender into adolescence while children from urban environments allowed for substantially more miscategorizations. This suggests that changes in the essentialist understanding or representation of gendered categories is not necessarily universal. That is, children from rural environments do not display the same decrease in essentialist reasoning about gender that children from urban environments do, though both start as essentialists. Thus, it must be considered whether or not children from urban environments really lose their essentialist bias or if it is simply hidden.

Indeed, Prentice and Miller (2005) showed that when male and female undergraduates were told that they had a particular perceptual style based on a performance to an irrelevant perceptual task, both men and women tended to say that members of their own gender were likely to share their perceptual style while members of the other gender were exceptionally unlikely to share their perceptual style. Such results demonstrate that even in the absence of cultural or stereotypical knowledge (i.e. when taught a new property that appeared to differ by gender), adult participants will extend properties along gender lines while refusing to extend it to the other gender category. Thus, once the new property appeared to be tied to the gender category, it similarly became tied to the essence, keeping it from being extended to those who did not have the appropriately gendered essence.

Other work by Haslam, Rothschild, and Ernst (2000) has shown that when asked to rate essentialist attributes of various social categories, undergraduates are willing to endorse such attributes for gender categories (e.g. that they are natural, immutable, and inherent). They do so equally for male and female categories on “natural kind”-like attributes (e.g. naturalness and
discreteness) while also extending more cohesive attributes to the category of female (e.g. informative and inherent). Similarly Bastian and Haslam (2006) showed that adult participants who indicated higher endorsement of various essentialist beliefs (e.g. immutability of categories or discreteness of categories) were more likely to endorse stereotypical gender traits. Thus, those adults who still endorse essentialist beliefs (e.g. innate potential, heritability, etc.) also still essentialize the category of gender.

Taken together, these results suggest a relatively strong essentialism for gender categories, especially among young children. Although the essentialism of gender categories by adults is less clear, there does seem to be an indication that gender essentialism is not completely absent in adults. Rather, there are hints from Rhodes and Gelman (2009) and Prentice and Miller (2005) that at least some older children and young adults are just as essentialist as younger children. Given that the vast majority of children start as essentialist and a subset retain their essentialist bias over time, we must consider the possibility that essentialism does not simply dissipate but that it is hidden by other cognitive processes (further discussion of this in the “Essentialism as a Type 1 Bias” section below).

**Religion / Religiosity**

Although less work has been conducted on the social category of religion, the study of this particular social category is especially compelling because, while gender and race have known biological and heritable bases, religion is a social category which individuals may objectively choose to enter and leave at any given time – though some religions like Judaism and Islam are also strongly intertwined with ethnic identity. Similarly, individuals may choose not to subscribe to any set of religious beliefs. Thus, while religion objectively lacks the biological bases of the above-outlined social categories, if people find it to be a meaningful, essentialized
social category, this would suggest that the essentialist bias for social categories is not simply biological in nature (Atran, 1990; Gelman & Hirschfeld, 1999) or at least does not necessarily refer to the actual biological basis of a category.

Studies on religion have mostly been conducted by Diesendruck and his colleagues in conjunction with their aforementioned work on ethnicity. For instance, Diesendruck and haLevi (2006) presented children with triads containing social category information and psychological trait information. Children consistently made inferences about property expression based on social category membership, including religious categories, as opposed to psychological traits even when there was no perceptual cue to religious affiliation (i.e. either an ambiguous face or only a label was presented). This suggests that religious labels denote the presence of a category essence allowing inferences to be made about members of different religious groups regardless of perceptual dissimilarity. However, as noted above, adults consistently made inferences based on psychological traits as opposed to social category membership. Again, this suggests two possibilities – either essentialism for religion is present in childhood but decreases and disappears in adulthood or that it is inhibited by other cognitive processes.

Birnbaum and colleagues (2010) also showed children several triads with individuals of differing religiosity (i.e. religious or secular) and membership in other social categories (gender, social status, ethnicity). In these triads, children of both Jewish and Arab backgrounds did not significantly use religiosity as a point of inference when reasoning about novel properties, though they were able to accurately identify religious category membership based on the pictures provided in the triads. Although religiosity membership was not found to be a greater source of inference than the other categories presented within the triads, this does not necessarily mean that religiosity is not essentialized. Indeed, with religiosity pitted against other categories in the
triad task, it may simply be that religiosity is not as essentialized as say ethnicity (which was found to be a significant source of inference as outlined above) but that it may overall be relatively highly essentialized.

Evidence for such an account can also be found from Haslam, Rothschild and Ernst (2000). When asked to endorse essentialist properties for various religious categories, young adult participants rated the category of “Jew” as both exceptionally natural and exceptionally cohesive, while the religious category of “Catholic” was rated highly on cohesive attributes. This suggests that while essentialism of religion may be relatively less than other social categories, it may still be relatively more essentialized overall than say a group based on psychological characteristics. This suggests that there may be a continuum of essentialism such that some categories are viewed as more essentialized than others as opposed to a binary distinction of “essentialized” or “not essentialized.”

Further, Deeb and colleagues (2011) found that when children were able to ask questions in order to try and discern category membership of an ambiguous person, older Jewish children frequently inquired as to the person’s religious affiliation, though younger Jewish children and Arab children did not. However, this implies that at least for some people (in this case older Jewish children) religion may represent a category containing a significant amount of information or be a significant source of inference.

Taken together, the results suggest that, while not as highly essentialized as say gender or ethnicity, religion may also be an essentialized social category. Again, it should be noted that results tended to show religion as essentialized by children and not adults, and even then not by all children (however, see Haslam, Rothschild, & Ernst, 2000). Yet, the possibility that religion may be essentialized at all, might suggest that essentialism is promiscuously instantiated in the
social domain. Religion represents one of many social categories that is not necessarily tied to a true ontology or “carving nature at the joints.” That is, individuals of religious categories do not have inherent-from-birth religious properties that we use to discern them from members of other religious categories. Instead, we are able to choose our religious affiliation and through a series of activities come to be members of that religious category. Because of this ability to alter religious membership, we should presumably find religious category membership to provide us little information about individuals. If members of a particular religious group have no inherent or innate properties that distinguish them from other members of other religious groups and if they can leave their religious category at any time for another, how can knowing someone’s religious affiliation be useful for inference? Still, the essentialist bias leads us to believe that religion is a potentially informative social category. Some possible mechanisms to explain this are outlined below.

**Proposed Mechanisms of Social Essentialism**

Within the literature on social essentialism, there appear to be two main views which have appeared regarding the instantiation of essentialism in the social domain. These two views propose different ways of representing the essentialism of a social category – either through a unified set of attributes or via two interacting factors.

**Social Categories as Natural Kinds**

Outlined by Rothbart and Taylor (1992), this view purports that social categories which are essentialized by an individual are represented and used as natural kinds, irrespective of their true ontological status. That is, if a social category is considered to have an essence, then it will be viewed as a naturally occurring category as opposed to simply a social construction. Such a view is a direct extension of Gelman’s (2003) work on natural kinds into the social domain. It
makes no distinction between essentialized natural kinds and social kinds and instead only draws differences between essentialized and non-essentialized kinds (i.e. artifacts).

Some evidence for this view can be found in Rhodes and Gelman (2009) and Taylor, Rhodes, and Gelman’s (2009) work on the differences or lack thereof between essentialism of natural kinds and social kinds. Indeed, such data suggest that at a young age, children are promiscuous essentialists and essentialize natural and social kinds to the same degree. This would seemingly support Rothbart and Taylor’s (1992) assertion that social kinds are basically indiscriminable from natural kinds within the cognitive system.

However, the developmental data also suggest that over time, natural and social kinds diverge in the extent to which they are essentialized. Across development, the behavioral properties of social categories tend to be less essentialized (Taylor, 1996; Taylor et al, 2009), the boundaries of social categories become less discrete and exclusive (Rhodes & Gelman, 2009), and in some cases social categories are discarded as a basis for inference in favor of other categories like those based on psychological traits (Diesendruck & haLevi, 2006). Yet, the essentialism of natural kinds does not seem to decline in the same fashion. If it were true that social kinds were simply viewed as natural kinds, then presumably they should maintain similar degrees of essentialism at any given point. Given that this is not the case for adults, an explanation of this discrepancy must be proposed to maintain the natural kind view. That is, if there are no intervening factors which attenuate the essentialism of social categories in adults and instead essentialism of social kinds simply dissipates with age, then the natural kind view of social categories seems untenable given that natural kinds do not show the same marked decline in essentialism.
Naturalness and Cohesiveness

Other social essentialist theorists have proposed two distinct factors for understanding social essentialism. Under this framework, there are two independent factors which contribute to the essentialism of social categories. Based on Haslam et al’s (2000) findings, as well as theoretical work from other social psychologists, these factors have been proposed to be “natural kind-ness” and “entitativity” (further referred to as “cohesiveness”). “Natural kind”-like attributes are how natural a category may be viewed or how discrete its boundaries might be (Haslam et al, 2000). That is, “natural kind-ness” refers to the extent to which people believe that a social category is naturally occurring in the world (however objectively false) – harkening back to the notion that categories represent a “carving of nature at the joints.” The second factor, “cohesiveness,” which describes how group-like a category might be, has been proposed in an attempt to better capture the essentialism phenomenon in the social domain (Yzerbyt et al, 1997; 1998; 2001). Cohesiveness may best be thought of as the extent to which a group shares similar behaviors and properties. In a sense, it represents the belief that a group member are similar enough to make inferences about the actions and goals of group members. Haslam and colleagues (2000) also attach beliefs about inherence – the notion that a group has an underlying reality despite surface differences – to the concept of cohesiveness.

Although the concept of cohesiveness is relatively similar to the notion of inductive potential (or the extent to which we find a group meaningful for making inferences), several studies have suggested that cohesiveness and naturalness independently contribute to the essentialism of social categories. For instance, Haslam et al (2000) found that two factors emerged in their participants’ endorsements of essentialist properties of various social categories. One of these corresponded to “natural kind”-ness while the other appeared to correspond to
cohesiveness. Further, these two factors varied independently for social categories such that some categories loaded highly on a natural factor but not on the cohesive factor and vice-versa. However, these were orthogonal and not unidimensional as categories could similarly load high or low on both factors. Thus, if the essentialism of a social category is truly the result of a combination of natural kind-ness and cohesiveness, then relatively high or low loadings on either or both factors would contribute to degrees of essentialism for different categories. That is, while some categories could load very highly on both factors, others might load highly on one factor and rather low on another, or only moderately on both factors. Thus, those that load highly on both factors could be considered extremely essentialized while those that load lower on both factors could be considered relatively less essentialized.

Further work by Thakkar (2003) found a dissociation between perceived cohesiveness of a group and the similarity of its members when making favorability judgments about a group. That is, when induced to think about the group as highly cohesive via a description about uniformly-acting members, participants were more likely to make extreme favorability judgments (either positive or negative). Further, mediational analyses found that only participants’ self-reported cohesiveness ratings predicted this change and not the extent to which they judged group members to be similar. Thus, featural similarity of members to one another appears to be different from cohesiveness. This seemingly reinforces the notion that two independent factors are used to make inferences about social groups – one being cohesiveness and the other being the extent to which a group is like a natural kind.

**Implications for Social Categories**

At the least, based on the above theoretical views, we can begin with the notion that essentialized social categories exhibit some measure of natural kind-ness; this is supported by
comparative data on social and natural kinds. The research on the development of essentialism suggests that there are few differences in the degree of essentialism of natural and social kinds in children, and children are equally willing to over-extend their social categories in the same way that they do natural kind categories (e.g., young children believe equally in the inherent potential of animal and social categories). This has several important implications for the structure of social categories and how we use them to make inferences.

First, as in the previous discussion of natural kinds, it would suggest that the category boundaries of social kinds are relatively rigid and discrete. That is, if social categories are at least viewed as “carving nature at the joints,” then people should treat the categories as defining something real and different in the world. As outlined above, there is evidence of this is in children and even in some adults (Rhodes & Gelman, 2009). Further, these boundaries are relatively as rigid as natural kind boundaries.

Second, we should use social kinds as a point of inference relatively as much as we do for natural kinds. That is, if social categories are at least partially viewed as or attributed to be natural kind-like, they should be used to make inferences with relatively similar frequencies as actual natural kinds. Again, some evidence of this exists based on Taylor and colleagues’ (2009) work comparing social and natural kinds. At least young children are willing to make inferences based on the innate potential of gender categories to roughly the same extent as they do for natural kind categories for both behavioral and physical characteristics.

However, a divergence in the strength of essentialism in natural and social kinds begins to occur in adolescence and this cannot be ignored. If we consider that we are social beings who use social categories on a daily basis, belong to social categories ourselves, and interact with members of a variety of social categories, it may be that we can attribute such a divergence is not
simply a diminishing of essentialism in the mind. Perhaps our social expertise or social conventions about the use of social categories are such factors. If this is the case, then we might be able to find that when controlling for such factors, adults and adolescents may exhibit relatively the same strength of social essentialist reasoning as do children, and relatively the same strength of social essentialism that adults show for natural kinds. One useful way to consider this divergence is within a dual-processing framework wherein the essentialist bias would not simply dissipate but rather it is simply working beneath the surface, moderated by other factors. Dual processing theory and how it might be applied to essentialism and especially essentialism of social categories is outlined in the following section.

**Dual Processing Theory**

The relatively early emergence of essentialism has led many researchers to focus on the childhood consequences of psychological essentialism within the reasoning process. However, as the previously outlined research suggests, there are hints of essentialism in adulthood as well. Yet, this essentialism is seen less prominently when adults perform the same reasoning tasks as children (e.g. Taylor et al, 2009; Rhodes & Gelman, 2009; Diesendruck & haLevi, 2006), though questions which tap into the conceptual underpinnings of essentialism (e.g. Haslam et al, 2000) reveal an essentialist framework which is still intact. That is, if essentialism were truly gone or reduced to only a bias for biological properties as some of the developmental comparisons suggested, it would be unlikely that adult participants would endorse essentialist concepts and attributes for wide ranges of categories as Haslam and his colleagues found.

This begs the question – what happens to essentialism in adults? If, as Gelman (2003) asserts, essentialism is a pervasive, early-emergent developmental bias which has a substantial effect on our reasoning processes, then where does it go as we develop and age? Haslam’s work,
as well as work by others (e.g. Eidson & Coley, 2014; Hampton, Estes, & Simmons, 2007;
Rhodes & Gelman, 2009) have noted essentialist responding by adults under varied paradigms
(e.g. switched-at-birth tasks, transformation tasks, category boundary tasks) and with varied
category types (e.g. gender, animal categories, race, respectively). Yet, these essentialist patterns
of responding are often not as extreme as is seen in children. Indeed, many of the findings of
essentialism in adults suggest a selective attenuation of essentialist responding. Adults seem to
favor essentialism of physical or biological properties but not necessarily behavioral ones and
favor essentialism of animal categories over social categories. Further, some categories seem to
be more essentialized than others even with a domain. Haslam and his colleagues showed that
categories differed in their rated naturalness or cohesiveness with some categories like “female”
or “Jew” having relatively high ratings of both factors, while categories like “Easterner” or
“blue-collar” had relatively low ratings for both factors, suggesting that the former categories are
relatively more essentialized than the latter.

Given that children start as promiscuous essentialists across domains and across
properties it would seem plausible that essentialism could be considered a basic cognitive
reasoning bias. Yet, in adulthood we find relative differences in essentialism between (and
within) domains and between different property types, suggesting that something has changed
within the reasoning system or in how essentialism interacts within the reasoning system.
Perhaps, as we age, some other factors intervene to quash our essentialist reasoning for social
categories or for behavioral properties. This is strongly consistent with the notion of a dual-
processing framework, wherein basic automatic biases may be intervened upon by effortful
reasoning to produce attenuated reasoning output. If this is the case, then we may consider
essentialism to be an early emerging, fast, automatic bias which is never replaced or disappears but which simply goes into hiding, buried under other reasoning processes.

As such, the section below will outline the theoretical considerations of dual processing theories. While there are many variations of dual processing theories proposed in multiple areas of psychology, the focus of the following section will focus on the theoretical similarities between these theories. That is, it will consider generally what is meant by dual processes and what these might look like generally in the reasoning process, without focusing too much on specific content areas initially.

**Dual-Processing Theory**

Dual-processing theories, in general, propose that two separate reasoning processes exist to influence reasoning and decision making. The difference between these two systems may best be thought of as the level of conscious access and influence which we have over those systems. Kahneman (2003) specifically cites the issue of accessibility as one of the core reasons for proposing dual processing theories. He notes that intuition often flies in the face of explicitly held knowledge (Tversky & Kahneman, 1971) with even experts being subject to their heuristic biases. Yet, we obviously are not subject to only our intuitions. We use a vast amount of information and knowledge to also effectively guide our decision making. As such, Kahneman proposed that the real distinction to be made in dual processing theories is between intuitive or automatic reasoning and explicit reasoning based on our large knowledge base.

Specifically, dual-processing theories generally classify fast and unconscious reasoning processes as Type 1, while Type 2 processes are generally thought of as slow, deliberate, and effortful (see Evans, 2003; Stanovich & West, 2000). While many variations exist for both the terms used to indentify Type 1 and Type 2, as well as the content associated with them (see
Evans, 2008 for a detailed comparison), they are at their heart meant to capture the unconscious and conscious reasoning processes. Thus as Evans and Over (1996) propose, our unconscious biases maintained in Type 1 may affect our behavior and reasoning unconsciously while Type 2 moderates the effects of those unconscious biases through the use of explicit reasoning and knowledge.

The fast, automatic nature of Type 1 reasoning has several implications for our reasoning processes. For instance, framing effects noted by Shafir (1993), McNeil and colleagues (1982) and Tversky and Kahneman (1981) suggests that beliefs about the statistical likelihood of success and even framing a problem in terms of success or failure has a strong influence on people’s reasoning outcomes and willingness to accept risk. That is, certainty of positive outcomes are generally highly favored, while certainty of negative outcomes are generally highly aversive. Yet, simple changes in instructions to emphasize or de-emphasize certainty while still providing the same outcome substantially changes participants’ responding. This emphasis on perceived certainty of outcomes, pushes Type 1 to different outcomes. That is, we may have a Type 1 bias to gain resources but with minimal loss. Thus, emphasizing uncertainty even with a similar statistical likelihood of outcome, shifts the output of Type 1. Yet, if Type 2 processes were engaged, we would be able to discern that regardless of the extent that certainty or uncertainty are emphasized, there is a set statistical likelihood of success which we can weigh for its potential return. Thus, as Kahneman (2003) points out, tapping into underlying beliefs about risk and certainty of outcomes via different stories or instructions automatically changes the framework within which people do computations or make decisions. Without active refutation of these beliefs by Type 2 processes, we are left to be subject to our underlying beliefs and intuitions.
Conversely, Type 2 reasoning may override our Type 1 intuitions through slow, effortful processing. Kahneman (2003) and others (e.g. Gilbert, 2002; Stanovich & West, 2000, 2002) suggest that Type 2 actively monitors Type 1 and intervenes when possible and appropriate. For instance, increasing the amount of effort required by a task may activate Type 2 reasoning (Kahneman, 2011) as long as the amount of effort required does not exceed tolerable limits. Thus, increasing difficulty by say asking a participant to count all of the f’s on a page (as opposed to identifying if an f is present at all) may activate Type 2 reasoning. Type 1 may have provided a general estimate of the number of f’s but by invoking the need for a specific count of items, Type 2 steps in to take control and attenuate the estimated response. However, when asking participants to determine the cost of items, Kahneman and Frederick (2002) found that participants were still highly susceptible to errors with many relying on relative magnitude to provide a close but reliably incorrect answer. Thus, errors may both be a failure of Type 1 to provide a correct response as well as a failure of Type 2 to correct the output of Type 1.

Further, limitations of Type 2 are especially noticeable when the system is placed under cognitive strain. Due to the slow, effortful nature of Type 2 processing, it is considered to be cognitively demanding. Thus, as in the above examples, though Type 2 is not always engaged, it can override Type 1 reasoning through explicit reasoning processes. Yet, engaging in such reasoning is not without cognitive cost. By limiting the availability of cognitive resources, the capacity for Type 2 to intervene on Type 1 output is decreased and it may not be engaged at all. Such interference may occur in the form of time pressure which will limit the temporal distance between the fast generation of output from Type 1 and responding, decreasing the likelihood that it can be intervened upon by Type 2 processes (e.g. Eidson & Coley, 2014; Finucane et al, 2000; Kelemen & Rossett, 2009). Similarly, dual-task interference (e.g. Gilbert, 1989, 1991) limits the
total number of cognitive resources available by splitting them between two different and demanding tasks. This limits the amount of cognitive energy that can be expended by Type 2 to moderate Type 1 output in a given task. Even conflict between preferred states and testing states (e.g. Bodenhausen, 1990) can reduce Type 2 involvement. Accordingly, those with higher working memory and intelligence (Stanovich & West, 2002) and “need for cognition” (Shafir & LeBoeuf, 2002) tend to be less susceptible to the interferences outlined above, presumably due to more efficient cognitive systems which better accommodate such interference, allowing Type 2 reasoning processes to be engaged.

**Dual-Processing Theories of Social Cognition**

Although essentialism may framed in general dual-processing terms, the most relevant specific dual-processing theory to the current work on social categories is targeted at understanding and reasoning about social groups. While dual-processing theories have been used to explain a wide variety of reasoning phenomena in various domains (see Evans, 2003, 2008), a dual processing theory of social cognition like that of Smith and DeCoster (2000) specifically aims to understand reasoning about social groups.

Under Smith and DeCoster’s social cognition dual-processing theory, two memory systems interact to influence our reasoning processes. In this model, Type 1 specifically handles the general acquisition of heuristic information. That is, knowledge about general regularities of groups (analogous to the categorical information conferred by belief in an essence) are stored and easily and quickly accessed by this Type 1. Thus, group generalizations (or essentialization) may best be thought of as a function of Type 1. However, Type 2 is concerned with rule-based reasoning. Thus, slow, effortful considerations of group dynamic or features may be handled by Type 2. However, it is important to note that Smith and DeCoster do not limit rule-based
reasoning to solely Type 2. Instead, they would propose that rule-based reasoning relies both on the generalizations made by Type 1 and the rule-based output of Type 2. However, generalizations need not interact at all with the rule-based reasoning of Type 2, indicating a one-way interaction from Type 1 to Type 2, but not in the reverse direction.

As Smith and DeCoster note, this interaction allows people to hold both generalizations (whether true or false) as well as personally conflicting beliefs. For instance, people may endorse typically negative stereotypes about a group while simultaneously maintaining beliefs that stereotypes are wrong (e.g. Devine, 1989; Plant & Devine, 1998). Thus, Type 1 contains associative stereotypical beliefs, while Type 2 may have rules like “Do not use a stereotype for group X” that may interact with the associative output of Type 1 to attenuate reasoning. This is not unlike the differences that we see between adults’ and children’s reasoning about social categories. That is, children start by making generalizations about social category members for various types of properties. Further, they appear to do this promiscuously for various types of social groups. Yet, over time we see a decrease in essentialist responding for social categories generally, and for behavioral properties specifically. Thus, a dual-processing theory like that of Smith and Decoster might suggest that children have a rampant Type 1 bias driving their categorization and inferences which is later attenuated in adults by Type 2 processes, including social conventions like “Don’t make generalizations about group X” or “Group X actually varies, and thus it is not appropriate to use past stereotypes.”

**Essentialism as a Type 1 bias**

As outlined above, within dual-processing frameworks, Type 1 biases are those which are fast, automatic, and belief-driven. As such, they provide the first output of the reasoning process based on limited external information and our own intuitions. However, this output may be
moderated and attenuated by Type 2 processes which rely on explicit knowledge systems and require slow, effortful processing. Given this description of the interplay between Type 1 and Type 2 processes, it follows that when individuals lack overlaying explicit knowledge to tap into (or those knowledge systems are not activated), Type 1 biases will prevail and drive the output of the reasoning system.

When Type 1 biases are considered in this manner, it would seem plausible that when individual’s lack information to attenuate possibly misleading Type 1 biases, these biases should be more prominent and more easily observable. When we consider the case of development, we certainly see that there are early, automatic biases that affect children’s reasoning including teleological reasoning where children ascribe goals and motivations even to inanimate objects such as rocks (Kelemen, 1999a, b) and psychological essentialism. As outlined above, essentialist reasoning is especially rampant in children for both natural kinds(e.g. Gelman, 2003; Heyman & Gelman, 2000a,b; Gelman & Markman, 1987) and for social kinds kinds (e.g. Diesendruck & haLevi, 2006; Hirschfeld, 1996; Taylor, 1996; Taylor et al, 2009).

In the case of both teleological and essentialist reasoning, children’s biases are initially promiscuous and are applied broadly to types of categories and for various types of properties. Thus, these biases may be thought of as unchecked by Type 2 processes resulting in overextended, incorrect inferences or the ignoring other information. This maps on very clearly to the idea of a Type 1 bias. That is, essentialism and other reasoning biases which are promiscuously applied, resulting in incorrect inference or ignoring other factors like perceptual similarity or existing differences between premises, may best be thought of as Type 1 biases which are not intervened upon by explicit Type 2 knowledge. However, once individuals come to know more about groups - how they function and how they vary – this promiscuity may
decrease with such biases either being only selectively applied or being attenuated by explicit knowledge stored in and used by Type 2 or potentially overwhelmed altogether as Type 2 completely overrides the erroneous output of Type 1.

Indeed, this is exactly the pattern that we see with psychological essentialism over the course of development. As Rhodes and Gelman (2009), Taylor et al (2009), and others have shown, from a very early age, children make essentialist responses about both natural and social kinds for behavioral and biological properties thus showing the promiscuous responding that one would expect from an unchecked Type 1 bias. However, as children age and presumably come to know more about categories and how their members function, these essentialist responses decline or become selective. For instance, as Taylor et al (2009) shows, essentialist responding with respect to natural kinds does decrease slightly as children age while essentialist responding about social kinds decreases dramatically, especially for behavioral properties. However, even into adulthood, people maintain many misconceptions about group variability and origins of animals which may allow the essentialist bias for natural kinds to only be slightly attenuated by Type 2 knowledge (e.g. Shtulman & Schulz, 2008; Shtulman & Valcarcel, 2012). Yet, we spend a vast majority of our lives as part of a social structure interacting with and being part of social groups. Thus, it is not surprising to see such a marked decrease in essentialist responding with respect to social kinds as we are more likely to acquire and use attenuating Type 2 information that would decrease reliance on the essentialist Type 1 bias.

If essentialism is indeed to be thought of a Type 1 bias which is intervened upon by Type 2 explicit knowledge as outlined above, this should mean that in cases where essentialism has been previously observed (e.g. essentialism of social kinds in children but not necessarily in adults) it should “re-emerge” or be left unchallenged by Type 2 processes when access to those
processes is limited. Thus, since essentialist responding about social kinds has been strongly noted and observed in children, though it is much more difficult to find in adults, the dual-processing framework should predict when and how essentialism of social kinds can be observed in adults.

As previously noted in the general review of dual-processing theories, limitations on the effectiveness of Type 2 processes may be imposed via time pressure, increased cognitive load through dual-task processing, or working memory restrictions (either experimentally imposed or as individual differences), resulting in increased rates of Type 1 output. Indeed, some evidence for such limitations allowing social essentialism to “re-emerge” has been found using time pressure. Eidson and Coley (2014) showed that undergraduates who were asked to perform a switched-at-birth task under time pressure produced more essentialist responses about gendered behaviors than those participants who completed the task under a time delay (though both groups gave near-ceiling essentialist responses about gendered biological properties, replicating the selective pattern seen by Taylor et al, 2009). Thus, those participants who did not have time to access explicit Type 2 knowledge systems gave more essentialist responses about the social category of gender. Further, under time pressure a significant portion of participants gave nearly complete essentialist responses, mirroring the responding typically seen in children.

Such evidence provides preliminary support for the notion that essentialism may act as a domain-general, fast, automatic Type 1 reasoning bias. Time pressure decreases reliance on and access to Type 2 processes, allowing Type 1 to produce essentialist output about social categories without being attenuated. However, further evidence is necessary to establish that the essentialist bias should be considered within the dual-processing framework.
Specific Aims of the Current Work

The current work explored the hypothesis that essentialism for social kinds is not merely pervasive in children (Gelman, 2003), but rather a persistent, fast, and automatic bias that remains in adults, which may be monitored and suppressed by T2 processes. If so, then essentialism might “re-emerge” when access to T2 processes is limited.

Specific Aim 1: Test the dual processing claim that essentialism can be thought of as a T1 bias.

Eidson and Coley (2014) provided initial evidence that time pressure increases essentialist responding about gendered behaviors. The claim that essentialism is a fast, automatic T1 bias can be examined in multiple ways. For instance, when essentialist responses are given at all, we might expect that they will be given faster than non-essentialist ones (Experiments 1, 4, and 5). Further, as Evans (2008) notes, while time pressure is one way in which to restrict T2 processes (which is expanded in Experiments 2 and 3) another method involves cognitive load (Experiment 1). Taken together, these pieces of evidence can bolster the claim that essentialism is indeed a T1 bias.

Specific Aim 2: Ascertain what aspects of T2 processing may be involved in limiting social essentialism.

Another key question is exactly what types of T2 processes are suppressing the essentialist bias in reasoning about social kinds. As noted by Eidson and Coley, there are at least two broad classes of possible factors – cognitive factors (e.g., belief in category variability or executive function) and social pragmatics (e.g., willingness to avoid prejudice). By measuring individual differences on factors which may influence social reasoning and their relationship to
essentialist thinking, we may be able to further specify what constitutes T2 editing of essentialist responses. These are examined in Experiments 2 through 5.

**Specific Aim 3: Determine the extent of adults’ essentialist reasoning.**

Given that essentialism and essentialist responding are comprised of multiple phenomena, this work will seek to obtain additional evidence of social essentialism by examining essentialist responding in other types of essentialist tasks. For instance, we employed the use of a transformation task (e.g. Keil, 1989; Rips 1989) and an inductive inference task (e.g., Gelman & Markman, 1986) to examine the extent of social reasoning about a variety of social categories. SAB tasks are more fully examined in Experiments 1 through 3 while Experiment 4 examines social transformations and Experiment 5 examines social induction.

**Experiment 1**

The aim of experiment 1 was to further test the dual processing interpretation of essentialist responding proposed by Eidson and Coley (2014) by manipulating cognitive load rather than time pressure. As Evans (2008) notes, while time pressure is one way in which to restrict the more explicit reasoning processes of Type 2 which can edit the fast, automatic responses of Type 1 another method of restricting Type 2 reasoning involves cognitive load. While time pressure restricts access to T2 processes by denying individuals time to adequately access or process information that could suppress T1 output – in this case essentialist responding – cognitive load accomplishes the same goal by recruiting available cognitive resources for other tasks thereby making them unavailable for use by T2 editing mechanisms.

Notably, De Neys (2006) pioneered the use of experimentally manipulated cognitive load to influence decision making on a syllogistic reasoning task. De Neys had participants view and memorize a dot pattern in a spatial matrix before completing the syllogistic reasoning task.
Participants were told that after completing the reasoning task, they would be required to recall the appropriate dot pattern. During the intervening time, participants were asked to evaluate the logical validity of syllogisms, including syllogisms which were logically valid but unbelievable or vice versa (e.g. “All mammals can walk. Whales are mammals. Therefore, whales can walk.” – logically valid but not believable). De Neys found that all participants, regardless of individual working memory capacity, were affected by cognitive load such that they were more likely to exhibit belief-based biases.

Although Eidson and Coley’s time pressure results offer initial evidence for the dual processing account of essentialism, validating those results with a cognitive load paradigm would provide additional evidence suggesting that essentialist reasoning is a result of Type 1 processes which may be attenuated through more explicit reasoning processes. To assess this account, a cognitive load paradigm was adapted from Goldinger and colleagues (2003) involving holding several non-words in memory while completing the exact same switched-at-birth task used in Eidson and Coley’s initial work. Of interest was whether the same increases in essentialist responding for gendered behaviors that were observed under time pressure could be observed under cognitive load. In this case, we would expect that performing the switched-at-birth task under cognitive load would be analogous to completing it under time pressure and that completing the task under no load would be analogous to completing it under a time delay. Should similar increases in essentialist responding be observed, we could conclude with increased confidence that once access to more explicit, effortful Type 2 reasoning processes has been strained or cut off, the fast, automatic essentialist reasoning of Type 1 begins to more strongly influence reasoning outcomes.
This experiment also offered the first opportunity to examine the relationship between how quickly participants responded and their willingness to make essentialist inferences. Given that this paradigm did not require participants to respond in a particular timeframe, we could examine a timing account of the dual processing framework. If essentialism is indeed a Type 1 process which is fast and automatic, we might expect that essentialist responses should be given faster than non-essentialist ones. It is also possible that this could interact with cognitive load in a speed-accuracy trade off such that cognitive load increases the difference in reaction time between essentialist and non-essentialist responses.

**Method**

**Participants**

Eighty-seven participants (44 in the cognitive loading condition) were recruited from Northeastern University’s undergraduate psychology pool.

**Design / Materials**

**Switched-at-Birth Task (SAB).** Following Eidson and Coley’s (2014) methodology, the same SAB task used by Taylor, Rhodes, and Gelman (2009) was administered to participants. Participants viewed four different vignettes – two in which a child (male or female) grew up with members of the same sex and, critically, two in which a child (male or female) grew up with members of the opposite sex. For instance, participants might read that Zillah, shortly after she was born, was sent to live on an island with all men. Participants would then be asked to decide what Zillah would be like once she was older. To this end, participants reasoned about both gendered behaviors (e.g., plays with tea sets or plays with toy trucks) and gendered physical properties (e.g., has girl blood or has boy blood). For each property pair, participants had to choose whether the child would have a stereotypically male or female property. In total,
participants received eight behavioral properties and eight physical properties in random order for each of the four vignettes (see Table 1).

**Cognitive Load.** Manipulated between subjects, participants were either assigned to complete the SAB task under cognitive load or under no load. Goldinger and colleagues’ cognitive loading paradigm was adapted for use with the SAB task. Participants were told that they would be completing a memory task in which they would initially be presented with a series of non-words which they would have to recall at a later time. Instructions emphasized that participants should do their best to retain as many words as possible for later recall.

Participants were always presented with six pronounceable non-words drawn randomly from a list of 55 non-words obtained from Rastle and colleagues’ (2002) ARC Non-word Database ([http://www.cogsci.mq.edu.au/~nwdb/nwdb.html](http://www.cogsci.mq.edu.au/~nwdb/nwdb.html)) before each SAB vignette. Each non-word was presented to participants for five seconds before moving to the next non-word. Critically, for participants under cognitive load, presentation of the non-words was followed by the SAB task (which was presented to participants as an irrelevant reading task), after which participants would recall as many non-words as possible from the initial set by typing them into the computer. However, participants who completed the SAB task under no load were asked to recall the list of non-words immediately after their presentation. Once recall was complete, participants would then read and respond to the vignette.

**Procedure**

Participants completed the experimental paradigm using a computer equipped with the SuperLab (Cedrus, 2006) stimulus presentation software. Instructions informed participants that they would be completing a memory task and would be presented with a list of non-words which they would then have to recall. They would also be completing an irrelevant reading task (the
SAB task). After being presented with an initial set of six non-words for five seconds each, participants in the no cognitive load condition would then immediately recall the words by typing them into an input box in the SuperLab (Cedrus, 2006) software before continuing on to a SAB vignette. Participants under cognitive load would withhold recall until after completing the intervening SAB task.

For the SAB task, participants would initially view one of the four vignettes regarding the target child (male or female) growing up on either a same- or opposite-sex island. After reading the vignette and indicating they were ready to proceed by pressing the spacebar, participants began to receive each of the 16 property pairs in random order. Participants would use the “f” key to indicate answers for items on the left side of the screen while they would use the “j” key to indicate answers for items on the right side of the screen with stereotypical male properties and female properties appearing equally on each side. Unlike Eidson and Coley’s time pressure manipulation, while reaction times for property pairs were recorded, they were in no way constrained; participants could answer as quickly or slowly as they saw fit. Participants would continue responding to each property pair until having answered all 16 items.

For participants in the cognitive load condition, the vignette would be followed up the unloading phase in which they would then recall their initial list of six non-words. Once both the vignette and recall were complete, participants would then receive one of the three remaining vignettes and repeat the process until the task was complete.

Results

Scoring

Participants responses were given a score of 1 if they gave an essentialist response (e.g. Zillah, a girl, would engage in a stereotypically female gendered behavior – “plays with tea sets”)
or vice versa for a male target) or a 0 if they gave a non-essentialist response (e.g. Zillah, a girl would engage in a stereotypically male gendered behavior – “plays with toy trucks” or vice versa for male target). Participants’ essentialist responses were summed for both behavioral and physical properties for each of the four SAB vignettes. Participants’ reaction times for each response were also recorded and averaged for behavioral responses and for physical responses. Further, reaction times were also split and averaged for essentialist and non-essentialist responses.

**Analyses**

**Cognitive Load.** As a manipulation check, participants’ average number of nonwords recalled for the switched vignettes were compared. An independent-samples t-test comparing average number of nonwords recalled in the each condition showed that participants in the no-loading condition ($M = 3.709, SD = .818$) recalled significantly more nonwords than those in the loading condition ($M = 3.148, SD = 1.049$), $t(85) = 2.780, p = .007, d = 0.596$. However, participants in both conditions recalled approximately half of the words, and mean recalls in both conditions were significantly greater than zero, $t_{\text{loading}}(44) = 19.910, p < .001$, $t_{\text{noloading}}(43) = 29.728, p < .001$.

**Essentialist Responses.** One of the main goals of experiment 1 was to determine whether Eidson and Coley’s time pressure results could be obtained using a cognitive loading paradigm. Specifically, does essentialist responding for gendered behaviors increase under cognitive load? If so, we would expect greater essentialist responding for gendered behaviors under cognitive load than under no load. To this end, a 2 (condition: load/no load) by 2 (target gender: male/female) by 2 (property: behavioral/physical) mixed measures ANOVA was conducted using condition as a between subjects variable, target gender and property as within subjects
variables, and essentialist responses in the switched scenarios as the dependent variable. Only a main effect of property was observed such that overall significantly more essentialist responses were given about physical properties ($M = 7.017$, $SD = 1.468$) than about behavioral properties ($M = 2.241$, $SD = 2.523$), $F(1,85) = 458.183, p < .001$. Critically, there was no main effect of condition and condition did not interact with property type, failing to find the same effect Eidson and Coley (2014) had observed using time pressure (see Figure 1).

**Response Times.** Although the expected property by cognitive load condition interaction was not observed, reaction times must be examined along with essentialist responding to provide a full account of participants’ reasoning processes. Under the essentialist dual processing account, if essentialism is indeed a Type 1 process, we might expect that essentialist responses should be generally given faster than non-essentialist ones. Because response times were not constrained, participants may also have circumvented the cognitive load manipulation by simply taking longer to respond to items for which they wished to make non-essentialist responses. That is, while the intent of the cognitive load manipulation was to limit access to effortful Type 2 reasoning processes and thus increase essentialist responding about gendered behaviors, because participants could take as long as they wanted to respond to any given item, this may have allowed participants to engage in a speed-response type trade-off such that they were still able to give non-essentialist gendered responses but at a cost of increased response time. If so, we would expect that response times for non-essentialist responses should be longer than for essentialist responses in the loading condition. Alternatively, although participants in the loading condition still recalled half of the nonwords on average, it may be that this was not sufficient load to require a speed-response trade-off. If this were the case, the dual processing account would still
ADULTS’ SOCIAL ESSENTIALISM

predict that essentialist responses when they are given may be faster in both conditions than non-essentialist responses.

To test these hypotheses, we examined the average response times for both physical and behavioral items and subdivided the response times for each by essentialist and non-essentialist responses. For instance, to the extent that an individual gave essentialist behavioral responses, the reaction times for those responses were averaged together and compared to the average response time for their non-essentialist behavioral responses. Because not all participants gave both essentialist and non-essentialist behavioral or physical responses, and participants who gave both an essentialist and non-essentialist response for one property type did not necessarily provide both responses for the other property type, analyses were conducted separately for behavioral and physical properties. To that end, a 2 (condition: load/no load) by 2 (response type: essentialist/non-essentialist) mixed ANOVA was conducted for each for each property type based on those individuals who had given at least one of each type of response (essentialist and non-essentialist) with condition as a between subjects variable and response type as a within subjects variable.

For behavioral properties, a marginally significant main effect of response type was observed such that overall participants took longer to give non-essentialist responses ($M = 3403ms, SD = 1876ms$) than essentialist responses ($M = 3006ms, SD = 1562ms$), $F(1,66) = 3.158, p = 0.08$. Similarly for physical properties, a significant main effect of response type was observed such that overall participants took longer to give non-essentialist responses ($M = 4219ms, SD = 2975ms$)$^1$ than essentialist responses ($M = 2712ms, SD = 1089ms$), $F(1,50) = 13.718, p = .001$ (see Figure 2). No main effects or interactions with cognitive loading condition

---

$^1$ Note that large differences and standard deviations in reaction time for physical responses are due to the small number of available non-essentialist responses to physical properties.
were observed for either property type suggesting that differences in response times are not attributable to the loading manipulation, but instead may reflect general differences in response times for essentialist and non-essentialist responses. The increased response times for non-essentialist responses are consistent with a dual processing account of essentialism wherein Type 2 reasoning processes requiring more effort and cognitive resources should take longer to produce non-essentialist responses than faster Type 1 essentialist reasoning output when it is allowed through.

**Discussion**

The original aim of experiment 1 was to replicate Eidson and Coley’s (2014) initial finding that essentialist responding about gendered behaviors could be increased under time pressure using the separate, but theoretically equivalent method of cognitive load. As our results showed, we were unable to replicate this finding in this particular cognitive loading paradigm. Indeed, essentialist responding about gendered behaviors in both the cognitive load and no load conditions was nearly identical to both Eidson and Coley’s delayed response condition as well as Taylor and colleague’s (2009) and Taylor’s (1996) findings that adults make very few essentialist responses about gendered behaviors at baseline. This at least suggests that there were no flaws with the SAB task itself.

This paradigm also offered the first opportunity to examine response times for essentialist and non-essentialist responses for behavioral and physical properties. Participants took longer to give both non-essentialist behavioral and physical responses. Again, this did not interact with loading condition, but it is consistent with a dual processing account of essentialism. That is, when participants provide essentialist responses, they provide them relatively quickly while non-essentialist responses take longer to produce. This is consistent with the hypothesis that
essentialist responses are the result of fast, automatic Type 1 processes, while non-essentialist responses require effortful consideration and intervening Type 2 processes.

Although we would expect that non-essentialist responses should overall take longer to give than essentialist responses, we might also expect that the discrepancy between the two would be even larger under cognitive load if participants circumvented effects of cognitive load on essentialist responses by simply taking longer to respond. Unfortunately, we did not find evidence to support this notion in these data. However, given that no main effects or interactions with cognitive load condition were seen for either the number of essentialist responses or the reaction times for essentialist and non-essentialist responses and that the average number of nonwords recalled in the loading condition was actually less than in the non-loading condition, we must seriously question whether participants in the loading condition were under sufficient cognitive strain. Given that the chosen cognitive loading manipulation involved pronounceable non-words, it may be that these were edged out in working memory by reading and conducting the SAB task, resulting in little to no load at all. A future version of this experiment should consider using a paradigm similar to De Neys (2006) in which a dot matrix of varying degrees of difficult was used to strain participants’ working memory.

Overall, although an exact replication of Eidson and Coley’s time pressure findings were not seen in Experiment 1 within the cognitive load paradigm, the ability to look at unconstrained response times for essentialist and non-essentialist responses provided a unique opportunity to assess a central timing claim of the dual processing account of essentialism. Indeed, we see that essentialist responses for both physical and behavioral properties are produced faster than their non-essentialist counterparts. This provides increased evidence for this account.
To further explore the dual processing account in accordance with Aim 1, Experiment 2 will return to the speeded SAB paradigm to determine whether participants placed under both delay and speeded conditions exhibit increases in essentialist responding. Several individual difference measures will also be employed to attempt to identify constructs involved in the suppression of essentialist reasoning in accordance with Aim 2.

**Experiment 2**

In accordance with Aims 1 and 3 outlined above, the experiment presented below was conducted to replicate Eidson and Coley’s (2014) previous findings that essentialist responding about gendered behavioral properties increased under time pressure. Eidson and Coley’s initial investigation used the same SAB task taken from Taylor et al (2009) outlined in Experiment 1 with half of participants completing the task under time pressure and half of participants under time delay. This allowed them to initially show that responding quickly to the switched-at-birth task produced a different essentialist response pattern from baseline levels of essentialism with participants under time pressure providing significantly more essentialist responses about gendered behaviors. This experiment expands on this methodology using a within-subjects time pressure manipulation; thus, a participant’s baseline delayed responding could be taken into account to determine whether an individual’s responding could be affected by time pressure. This allowed us to examine not only participants’ speeded essentialist responses, but also the degree to which time pressure affected responding compared to the delayed baseline. To this end, participants completed the same SAB used in experiment 1. As in Eidson and Coley (2014), a time pressure manipulation was used to elicit increased essentialist responding about gendered behaviors. Critically, participants completed both a speeded *and* delayed version of the SAB.
This enabled a more direct test of the claim that time pressure can increase essentialist responding by determining whether individual subjects were affected by time pressure.

Further, in accordance with Aim 3, we were interested in determining what factors might predict the change in essentialist responding under time pressure or changes in responding between the delay and speeded conditions. Of various individual differences which may be associated with essentialist responding or susceptibility to the time pressure manipulation, three groups of factors seemed most likely – conceptual factors, social factors, and executive function.

The first, a more conceptual factor, was the extent to which participants believed that categories actually varied. As previously discussed, variability within a category is counter to general essentialist reasoning. That is, if all category members have an essence and this essence gives rise to category-based properties, then presumably all category members should exhibit those category-based properties resulting in very little category variability (e.g., Rhodes, Gelman, & Brickman, 2008; Rhodes & Brickman, 2010; Shtulman & Shulz, 2008). Thus, individuals who perceive or understand categories as having variable properties might also hold less essentialist beliefs. Thus, those who have a greater belief in category variability might exhibit overall fewer speeded essentialist responses.

A second possible predictor of individual differences in essentialist responding may be various social-motivational factors such as the extent to which we are concerned about how our decisions and predictions about other people might be perceived. Plant and Devine (1998) and Dunton and Fazio (1997) and their colleagues have identified both an internal and external motivation to avoid appearing prejudiced against members of other social groups. While external motivation to avoid prejudice refers to societal norms or pressures to avoid making stereotypical judgments about others, internal motivation to avoid prejudice refers to an individual’s personal
beliefs that such judgments are wrong or should be avoided. Thus, individuals who feel highly motivated to avoid prejudice – either internally or externally – may actively suppress their essentialist tendencies showing associations with degree of change in essentialist responding. As with category knowledge, accessing such motivation presumably requires time and cognitive resources. That is, while their motivation to avoid prejudice may lead them to make relatively few if any essentialist judgments under a time delay, once access to that motivation is restricted by time pressure, essentialist responding may increase dramatically. This is in contrast to someone who has little motivation to avoid prejudice and thus would produce relatively consistent essentialist responses show less change under time pressure.

Another social individual difference which may be predictive of essentialist responding may be political affiliation. Political affiliation – especially conservative ideology – has previously been shown to be associated with stereotyping and prejudicial attitudes (e.g., Napier & Jost, 2008; Sidanius, Pratto, & Bobo, 1996). Given the proposed links between social essentialism and prejudice (e.g., Allport, 1954; Haslam, Rothschild, & Ernst, 2002) and the known association between conservative political affiliation and prejudicial attitudes, it is possible that those who adhere to more conservative ideology will provide more speeded essentialist responses overall compared to their liberal counterparts. Similarly, right-wing authoritarianism (RWA - Altemeyer, 1981; 1996; 2006) which may best be thought of as a personality measure which generically assesses not participants’ own authoritarian tendencies, but rather submission to and acceptance of authoritarian leadership, has been shown to be correlated with increased prejudice and stigma against minority groups, as well as with increased inferential errors (Altemeyer, 1996). While those high in RWA tend to exhibit conservative world-views (Altemeyer, 2006), RWA is only modestly correlated with conservative political
affiliation thus requiring independent measurement. We might expect that individuals with higher RWA, like those with more conservative political affiliation, would exhibit overall greater speeded essentialist responding.

Another socially influenced difference which might influence essentialist reasoning is self-reported gender role as measured by the Bem Sex Role Inventory (Bem, 1978). Work by Deaux (1985) and Bem (1981) suggests that gender role (i.e. how masculine or feminine someone is) may reflect culturally accepted attributes of “maleness” or “femaleness.” Given that we asked participants to reason about highly stereotypical male and female properties within the SAB task, participant gender and gender roles may also affect participants’ essentialist responding. For instance, Eidson and Coley (2014) found that male participants were significantly more essentialist than their female counterparts. These findings were consistent with previous research that male gender roles are generally viewed as less flexible (e.g., Berndt & Heller, 1986) and that male participants tend to allow less flexibility in gender roles (Levy, Taylor, & Gelman, 1995). Following this logic, we might expect individuals who either perceive themselves as more masculine or less feminine to provide more speeded essentialist responses overall analogous to the finding that males are generally more essentialist than females.

The third possible type of predictor to consider is executive function. That is, even if neither of the above factors were implicated in changes in essentialist responding our general inhibitory control system could also affect essentialist reasoning. In analogous work on teleological reasoning in adults, performance on the Stroop task, one measure of executive function, has been shown to predict production of teleological explanations when participants are placed under time pressure (Kelemen & Rosset, 2009). Given that teleological reasoning and essentialism are similar in nature as cognitive biases – both emerge early in development but
come more selective over time – we might expect executive function to similarly be predictive of changes in essentialist reasoning. That is, if psychological essentialism is indeed a Type 1 bias which is fast and automatic, but inhibited by various Type 2 processes (whether they are adequately captured by those outlined above or not), then presumably the extent to which we are good inhibitors at all should be predictive of the extent to which we exhibit changes in essentialist responding between the two timing conditions. Thus, good inhibitors might produce generally fewer speeded essentialist responses overall, but when the inhibitory mechanism is overwhelmed via time pressure, then we might see a marked increase in essentialist responding.

By taking these individual differences into account (see Table 2 for a summary of predictions for all measures), we may begin to accrue evidence not only that essentialist reasoning does “re-emerge” when access to Type 2 processes are limited but we may also begin to understand exactly which aspects of Type 2 processing are involved in inhibiting essentialist responding in the first place. If we assume that essentialism is indeed a fast automatic bias, then individual differences which either correlate with number of essentialist responses or degree of change in essentialist responding – either positively or negatively – can be inferred to either promote or inhibit essentialist reasoning generally. Given that essentialist responses about physical properties have been ubiquitous in previous work (e.g., Eidson & Coley, 2014; Taylor, Rhodes, and Gelman, 2009) and in Experiment 1 – individual differences that correlate with essentialist responses about behaviors will be most informative. Further, although associations with essentialist responding in both conditions would help to flesh out any associations observed, since essentialist responses about behaviors are exceedingly rare at baseline, correlations for speeded responses which more closely approximate adults’ underlying essentialist tendencies will be examined alone. Separately, the difference in essentialist responding about behaviors
between delay and speeded conditions can be conceptualized as the degree to which participants are engaged in active suppression of essentialist reasoning. Thus, individual differences which show associations with changes in responding represent either a resistance or susceptibility to increased essentialist responding under time pressure depending on the direction of the association.

Method

Participants

Eighty-seven adult participants were recruited from the Northeastern University undergraduate psychology pool. Participation fulfilled partial course credit. The study was approved by the Northeastern University IRB.

Materials and Design

Switched at Birth (SAB). The switched-at-birth task was identical to that of Experiment 1 with participants reasoning about both a male and female target and about behavioral and physical properties. Critically, participants received one vignette under speeded conditions, and one under delayed conditions.

Thus, independent variables in the SAB task were all manipulated within-subjects and included the gender of the target, the type of property, and the speeded or delay condition while the dependent variable was number of categorical or essentialist decisions made for each property.

Individual Difference Measures. Each individual difference measure is outlined below including cognitive measures (i.e., belief in variability), social measures (i.e., motivation to avoid prejudice, political affiliation, RWA, gender role), and executive function measures (i.e., Stroop and SIC).
Belief in Category Variability. Adapted from Shtulman & Shulz (2008), the category variability task asks participants to consider the prevalence of the same properties used in the SAB among males and females. For instance, participants were asked “What percentage of males do you think play with toy trucks?” Participants then gave a response ranging from 0 to 100 percent.

Participants were presented with 4 behaviors (2 male and 2 female), 4 preferences (2 male and 2 female), 4 physical properties (2 male and 2 female), and 4 category-based properties (2 male and 2 female) in random order and were asked the percentage of males and females that had that particular property or behavior. A variability score for each property was computed by subtracting the percentage of the opposite gender that had the property from the percentage of the expected gender that had the property (e.g. for “Play with toy trucks” the percentage of females who played with toy trucks was subtracted from the percentage of males who played with toy trucks in order to produce a variability score for this particular male behavior). Variability scores were then averaged across property type for each gender resulting in a single variability score for behavior and physical properties. Higher scores represent greater endorsement of sex-typed beliefs which is consistent with greater essentialist belief while scores closer to zero represent a lack of endorsement of sex-typed beliefs.

Motivation to Avoid Prejudice. Although originally developed to assess attitudes about racial prejudice, items from both Plant and Devine (1998) and Dunton and Fazio (1997)’s motivation scales were adapted to assess motivation to avoid prejudice more broadly for the purposes of the current work by removing references specific to race. Internal motivation items assess the extent to which participants are motivated by internal or personal beliefs to avoid
appearing prejudice while external motivation items assess the extent to which participants are motivated by external or social factors to avoid appearing prejudice.

Adapted items include the entirety of Plant and Devine’s internal and external motivation to avoid prejudice scale (e.g. “I attempt to avoid acting biased against members of other groups because it is personally important to me” or “I try to hide any negative thoughts about members of other groups in order to avoid negative reactions from others,” respectively) as well as items from the motivation to avoid prejudice factor of Dunton and Fazio’s scale (e.g. “It bothers me a great deal when I think I’ve offended someone, so I’m always careful to consider other people’s feelings”[internal] or “In today’s society it is important that one not be perceived as prejudiced in any manner”[external]). Participants assessed their agreement with these statements on a scale from 1 (strongly disagree) to 9 (strongly agree). Ratings were then averaged across items for both external and internal motivation statements.

**Political Affiliation.** Although many measures of political affiliation simply ask participants to identify as liberal or conservative (e.g. Jost, 2006; Jost, Federico, & Napier, 2009), this fails to capture the possible breadth of political ideology which may include both economic and social issues. As such, we used Pratto and colleagues (1994) method for assessing political affiliation by asking participants to rate how liberal / conservative they were about the following issues: social issues, economic issues, foreign policy issues, and how liberal or conservative they were generally. Each item was rated on a scale from 1 (extremely liberal) to 7 (extremely conservative). Ratings were averaged across all items in order obtain a single political affiliation measure.

**Right Wing Authoritarianism (RWA).** Participants completed Altemeyer’s (2006) RWA scale by rating their agreement with 28 separate statements regarding prejudicial attitudes (e.g.}
“There is absolutely nothing wrong with nudist camps.”) and trust in government and authority (e.g. “It is always better to trust the judgment of the proper authorities in government and religion than to listen to the noisy rabble-rousers in our society who are trying to create doubt in people’s minds.”). Ratings were made on a 1 (strongly disagree) to 9 (strongly agree) scale with half of items being reversed scored such that higher ratings indicated lower RWA beliefs. Responses were averaged across all items to produce a single RWA score.

**Bem Sex Role Inventory (BSRI).** The BRSI (Bem, 1978) contains 60 adjectives and phrases that are masculine (20 items), feminine (20 items), or neutral (20 items) in nature. Participants rated how much they personally identified with masculine (e.g. assertive, analytical), feminine (e.g. sensitive to the needs of others, flatterable), or neutral (e.g. talkative, friendly) items. Responses were made on a scale from 1 (never or almost never true about myself) to 7 (always or almost always true about myself). Critically, masculine and feminine items were averaged separately resulting in both a masculinity and a femininity score for each participant.

**Stroop.** Using the same general methodology developed by Stroop (1935), participants viewed three different blocks of stimuli: colored X’s, color words which are colored with congruent font (e.g. **BLUE**), or color words which are colored with an incongruent font (e.g. **RED**). Participants responded based on the color of the font, regardless of the word. Critically, in the incongruent or interference condition, participants must inhibit the impulse to respond based on the color name as opposed to the actual color of the font.

Participants first received a block of 36 trials containing colored X’s (e.g. **XXXX**). 9 trials were presented for each of four different colors: blue, red, yellow, and green. This was followed by a block of 36 trials containing color words presented in a matched color (e.g. **BLUE**). Again, 9 trials were presented for each of the four different colors. Lastly, a block of 36
trials was presented with color words presented in a mismatched color (e.g. RED). For each color word, three trials were presented for each of the mismatched colors (e.g. three trials of the word “RED” in the color blue, three in the color yellow, and three in the color green). In all cases, participants were instructed to respond with the color word while ignoring color of the font.

Number of errors and reaction times in the congruent condition were recorded and subtracted from those in the incongruent condition in order to account for participants’ baseline performance.

**Semantic Inhibitory Control (SIC).** The SIC task (,) was used as a second measure of executive function and inhibition. The task required participants to read a sentence such as “_____ is a color.” followed by presentation of a word or nonword that was related or not. For instance, participants could receive the word “Purple” (related word), “Plurple” (related nonword), “Trout” (unrelated word), or “Blicket” (unrelated nonword). Participants then made a decision about whether or not the presented word or nonword made sense in the previously seen sentence.

Critically, participants responses to related nonwords were examined against their responses to unrelated nonwords in order to determine level of inhibitory control. For instance, since “plurple” is not actually a word, participants should respond that it does not make sense in the sentence. However, a failure to inhibit “plurple’s” perceptual similarity to “purple” may lead a participant to respond that it is, indeed, a color. “Blicket” is easily identifiable as a nonword and perceptually dissimilar to other color words rendering it a suitable control for related nonwords, though the response to both “plurple” and “blicket” should be that it does not complete the sentence in a meaningful way. Altogether 12 of each word type were randomly presented with different sentences.
The number of errors made in the unrelated nonword condition was subtracted from the number of errors made in the related nonword condition to indicate levels of inhibitory control. The same differences were calculated for reaction times.

**Counterbalancing and Randomization.** Participants received all of the above assessments in one of eight counterbalanced orders. However, the tasks were not completely randomized due to experimental considerations. The switched-at-birth tasks and executive function measures were temporally spaced to avoid diminishing time pressure effects due to response familiarity and allowed participants time to recover from one executive function measure before completing the other. The variability measure was always presented last to avoid affecting essentialist responses in the SAB as both used the same properties. The counterbalanced orders followed this general format: an inhibitory control measure first, a switched at birth task where subjects reasoned about either a male or female target in either speeded or delayed conditions, randomized individual difference measures (Motivation, RWA, Political Affiliation, BSRI), the second switched at birth task in the second condition with subjects reasoning about the second gender category, the remaining inhibitory control measure, and beliefs about variability. Further, items were presented in random order within each task.

**Procedure**

**SAB.** The presentation of the SAB and timing conditions was identical to that of experiment 1 with participants critically completing both a speeded and delayed SAB.

**SIC.** The SIC was presented using SuperLab (Cedrus, 2006) software. Participants were given instructions to respond to items as written to ensure they did not simply believe that related nonwords were typos made by the experimenters. After reading instructions, participants first engaged in practice trials to familiarize themselves with the task. A sentence would appear on the
screen such as “________ is a vegetable.” Once participants pressed a key, the sentence disappeared and was replaced with either a related word (“carrot”), an unrelated word (“trout”), a related nonword (“broccoli”), or an unrelated nonword (“blicket”). Participants then responded whether the presented word made sense in the sentence using the “0” key to respond that the word made sense or the “3” key to respond that the word did not make sense. Participants were given feedback during the practice trials.

After completing the practice trials, participants continued to receive sentences like the one above followed by one of the four word types. After making their decision as to whether or not the word made sense in the presented sentence, participants saw a fixation cross for approximately one second before the next sentence appeared on the screen. Sentences-word pairs were presented randomly until all 48 pairs had been completed.

**Stroop.** The Stroop task was presented using SuperLab (Cedrus, 2006) software. After reading instructions, participants first engaged in practice trials to familiarize themselves with the task. A word or series of letters was presented on the screen in a given color (e.g. **BLUE**). Participants were told to press the key corresponding to the color in which the word was presented, regardless of the word itself. Participants pressed “q” for red, “w” for yellow, “o” for green, or “p” for blue. During the practice trials, participants were told whether or not their choice was correct in order to ensure that participants understood the nature of the task.

After completing the practice trials, participants continued to see words appear on the screen in three different blocks – a color only block (where participants saw colored Xs - **XXXX**), a color-match block (where participants saw a color word in its appropriate color - **BLUE**), and a color-mismatch block (where participants saw a color word in an unmatched color - **BLUE**). Participants pressed the appropriate key for the color of the word. After a key press,
participants would be presented with a fixation cross for 1 second before receiving the next word. Trials continued until participants had completed the experiment.

**Individual Difference Measures.** All individual difference measures and items within each measure were presented in random order using the online survey software Qualtrics (Qualtrics, 2013). Instructions were presented before each individual difference measure. Participants moved through questions at their own pace by clicking the appropriate bubble corresponding to their answer before clicking the “next” button at the bottom of the screen.

**Results**

**Scoring**

**SAB.** For switched at birth tasks, participants were scored as having given an essentialist response if the category chosen matched the described child’s birth category. The total number of essentialist responses was summed for both physical and behavioral properties in each of the timing conditions.

**Individual Difference Measures.** All individual difference measures were scored in accordance with their individual conventions. For survey measures, this involved averaging participants’ ratings to critical items to produce a single score for a given measure or construct (e.g., ratings for the motivation to avoid prejudice questions were averaged separately for internal and external motivation items to produce two separate average scores). For the executive function measures, this involved counting the number of errors made in critical conditions and subtracting errors in baseline conditions as well as recording the reaction times for correct responses.

**Analyses**
Essentialist Responses. Essentialist responding was conducted both for participants’ average essentialist responding and individual response patterns. We expected that essentialist responding about behaviors should show increases in both cases.

Mean Essentialist Responses. The first goal of this experiment was to attempt to replicate Eidson and Coley’s finding that time pressure increases categorical responding about gendered behaviors using a within-subjects manipulation of time pressure. To test this, we conducted a 2 (condition: delay, speeded) by 2 (property: behavioral, physical) by 2 (participant gender: male, female) repeated measures ANOVA on the frequency of categorical choices. Critically for our hypothesis, the ANOVA revealed a property by condition interaction \((F(1,85) = 50.439, p < .001)\) (see Figure 3).

Paired-sample t-tests showed that participants made significantly more category-based responses about gendered behaviors under time pressure \((M = 3.72, SD = 2.78)\) than under a time delay \((M = 2.21, SD = 2.49)\), \(t(86) = 5.48, p < .001\), and that participants made significantly fewer category-based responses about physical properties under time pressure \((M = 6.39, SD = 2.06)\) than under a time delay \((M = 7.00, SD = 1.44)\), \(t(86) = 3.65, p = .001\). Thus, we replicated Eidson and Coley’s initial finding that time pressure increased essentialist responding about gendered behaviors even when time pressure is manipulated within subjects.

Results also showed a main effect of property such that overall participants made significantly more categorical responses about physical properties \((M = 6.70)\) than behavioral properties \((M = 2.97)\) \((F(1,85) = 269.677, p < .001)\), and a main effect of condition such that overall participants made significantly more categorical responses under time pressure \((M = 5.06)\) than under a time delay \((M = 4.60)\) \((F(1,85) = 6.894, p = .010)\). However, participant
gender did not produce a significant main effect or higher order interaction with property or condition.

**Individual Response Patterns.** Overall, the group results suggest that participants’ essentialist responding about gendered behaviors increases under time pressure. Another way to examine this effect is through individual response patterns to determine if a significant number of individual participants changed their response patterns from delay to speeded conditions. Although analyses of individual response patterns in previous work has employed a two-way Chi-Square (e.g., Eidson and Coley, 2014), due to the repeated measures time pressure manipulation that was not possible. Instead a McNemer’s Chi-Square test for paired data was used. This requires that participants be typed into two nominal variables for two repeated measures. In this case, participants were typed as essentialists if they gave 6 or more essentialist responses about behaviors out of 8 or non-essentialist if they gave less than 6 essentialist responses for both speeded and delay conditions. Counts were then submitted to a McNemer’s Chi-Square test to determine if significantly more participants exhibited consistent essentialist response patterns under speeded conditions \( (N = 25) \) than delay conditions \( (N = 11) \) (see Figure 4). Results showed that indeed significantly different patterns of responses were obtained under speeded conditions than delayed conditions with more participants exhibiting consistent essentialist response patterns about behaviors under time pressure, \( p < .001 \). This suggests that not only does essentialist responding under time pressure increase in general, but that individuals are also more likely to give a set of consistent essentialist responses under time pressure compared to a delay.

**Individual Differences.** The second goal of this experiment was to determine what factors might predict essentialist responding about gendered behaviors and the degree of change
in essentialist responding between delay and speeded conditions. Regression analyses were performed on the number of essentialist responses about behaviors under speeded conditions as well as the degree of change in essentialist responding between the two conditions. To assess the degree of change we calculated a difference score for each participant by subtracting the number of essentialist responses made under time delay from the number of essentialist responses made under time pressure. For speeded essentialist responses, various individual differences which may facilitate or inhibit essentialism generally may generally predict essentialist responding under speeded conditions – either positively or negatively. To the extent that any individual difference may be part of a Type 2 mechanism suppressing essentialist Type 1 responding, we might expect that responses to a given measure would be especially predictive of the change in essentialist responding under different timing conditions.

All individual difference measures were entered as independent predictors into separate multiple regression models with change in essentialist responding about behaviors or speeded behavioral responses as the dependent variable. Given the aforementioned possibility of effects of participant gender, the same set of regressions was also performed separately for both male and female participants. Due to concerns about the number of predictors used in the model, we conservatively only consider effects of individual predictors with $p < .05$ as significant and $p < .06$ as marginally significant. All regression coefficients and significance levels are summarized in Table 3 and outlined in the following subsections.

**Models.** Overall, regression models were generally poor at predicting speeded essentialist responding about behaviors and degree of change from delay to speeded conditions. However, taken together, individual differences were predictive of male participants’ degree of change in essentialist responding, $F(12,31) = 12.186, p = .040, R^2 = .458$. More importantly, specific
individual difference measures which were significantly predictive of essentialist responding are outlined below.

**Belief in Category Variability.** Although not predictive of essentialist responding or degree of change overall, greater sex-typing of gendered behaviors was found to be marginally predictive of male participants’ degree of change in responding from delay to speeded conditions, $\beta = .331, p = .058$. This suggests that males who perceive certain behaviors to be more aligned with one gender exhibited larger degrees of change in essentialist responding. Thus, they exhibited relatively fewer essentialist responses about gendered behaviors under time delay and relatively more essentialist responses under speeded conditions. Conversely, male participants who do not necessarily sex-type behaviors are more consistent in responding between the two timing conditions. Thus, we did not find evidence for our original prediction that those who perceive more variability among gendered categories (and less sex-typing) would exhibit less essentialist responding overall, but sex-typing may be related to suppression of essentialist responding.

Variability scores for participants’ beliefs about distribution of physical properties was not significantly predictive of either speeded responding or degree of change in any instance.

**Motivation to Avoid Prejudice (Internal / External).** Neither internal or external motivation to avoid prejudice was a significant predictor for participants generally of either speeded responses or degree of change. However, for male participants, internal motivation to avoid prejudice was marginally associated with degree of change in essentialist responding, $\beta = -.354, p = .087$. This suggests that those who rate themselves as having greater personal beliefs that stereotyping should be avoided exhibit more consistent essentialist responding in the two conditions while those who feel less internal pressure to avoid stereotyping actually exhibit
greater responding in the speeded condition compared to the delayed condition. This is counter to our original hypothesis that those who exhibit motivation to avoid prejudice would show increases in essentialist responding under speeded conditions compared to a delay.

**Political Affiliation.** Political affiliation similarly did not predict participants’ overall speeded responses or degree of change. Interestingly, however, for female participants political affiliation significantly predicted speeded essentialist responses about gendered behaviors (β = .551, p = .020) such that females who rated themselves as more conservative exhibited greater speeded essentialist responding. This was consistent with our prediction that conservative ideology would be associated with increased essentialism. However, the opposite was true for male participants. Political affiliation was predictive of both speeded essentialist responding (β = -.457, p = .007) and degree of change (β = -.517, p = .002) though in both cases males who rated themselves as more liberal exhibited greater essentialist responding and greater degree of change. While the latter result suggests that political affiliation may also be implicated in suppression of essentialism, the former finding is both in the opposite direction of female participants’ relationship and counter to our original hypothesis that those with more conservative ideology would experience greater speeded responding overall.

**RWA.** RWA was not predictive of participants’ overall speeded essentialist responding or degree of change, nor was it predictive for male and female participants separately.

**Gender Role (Masculinity / Femininity).** While femininity was not predictive in any instance, masculinity was predictive for both participants’ overall speeded responses (β = .242, p = .057) and for male participants’ degree of change in essentialist responding, β = .349, p = .057. The former result is consistent with our prediction that to the extent that masculinity can be conceptualized as culturally stereotypical attributes of “maleness” more masculine individuals
should exhibit greater essentialist responding. However, the finding that masculinity is also associated with increased change in essentialist responding between delay and speeded conditions suggests that gender role may also be implicated in suppression of essentialist responding.

*Executive Function (Stroop / SIC).* Executive function measures – both Stroop and SIC – showed no predictive relationship with participants overall essentialist responding or degree of change nor for male and female participants separately.

**Conclusion.** These results suggest that both the change in essentialist responding and speeded responding may be predicted by both conceptual (e.g., beliefs about variability) and social (e.g., political affiliation) factors. However, the results also show two opposite patterns of these factors for male and female participants. This suggests that while essentialist responding may be predicted, there is no easy universal one-to-one predictor of such responding.

**Discussion**

The goal of Experiment 2 was to both replicate and extend past findings that essentialist responding about gendered behaviors increased under time pressure, providing evidence that when Type 2 processes were limited, essentialism would “re-emerge” in the reasoning system, while also attempting to determine what factors might be responsible for inhibiting essentialist responding at baseline.

The results outlined above suggest that even when using a within-subjects time pressure manipulation such that all participants received the SAB task under both time delay and time pressure, a significant increase in essentialist responding is observed for gendered behaviors when participants are under time pressure. This suggests not only that the finding is robust, but it provides further evidence that individual participants’ responses were more essentialist under...
time pressure than under a time delay. This is consistent with the idea that when access to Type 2 reasoning is limited, essentialism is “re-emerging” within the reasoning system producing greater essentialist responding under time pressure.

Interestingly, a significant decrease was observed for essentialist responding about physical properties under time pressure. Although previous work has not found such a decrease in essentialist responding, this may simply be due to the increased power of a repeated measures design and analyses. Subsequent testing will be necessary to determine if this finding will be replicable.

For the individual difference measures, most measures were predicative of essentialist responding or degree of change between timing conditions in some instance. Belief in variability – as measured by participants’ willingness to sex-type behavioral properties – was predictive of the degree of change in males responding between timing conditions. Thus, we found no direct evidence that belief in variability was associated with overall essentialist responding. Instead, it was implicated in the suppression of essentialist responding. Interestingly, those who were more willing to sex-type exhibit greater speeded essentialist responding than under a delay which is somewhat consistent with the notion that sex-typing would be associated with speeded essentialist responding. Thus, it may be that under time pressure, essentialist responding is coming into line with beliefs.

Although external motivation to avoid prejudice was not associated with either speeded essentialist responding or degree of change in responding, internal motivation to avoid prejudice was negatively associated with increased degree of change in essentialist responding among male participants. That is, males who reported less internal pressure to avoid stereotyping exhibited greater change in essentialist responding from time delay to speeded conditions. This is counter
to our original predictions based on Plant and Devine’s (1998) previous work. We would have expected that those with pressure to avoid stereotyping would give relatively few essentialist responses at baseline and then once access to those motivations were restricted through time pressure, essentialist responding would comparatively increase. This suggests either that motivation to avoid pressure simply does not suppress essentialism as we may have thought or even promotes essentialist responding under time pressure. Alternatively, it may also be that, at least for males, essentialist beliefs come into line with their lack of motivation to avoid prejudice under time pressure.

Political affiliation also showed significant associations with essentialist responding and degree of change. Interestingly, for female participants conservative political affiliation was associated with increased speeded essentialist responding. This is consistent with past findings (e.g., Napier & Jost, 2008; Sidanius, Pratto, & Bobo, 1996) and our prediction that conservative ideology would be associated with increased essentialist responding in general. However, the opposite was true for male participants. Males who self-reported as more liberal exhibited increased essentialist responding under time pressure. While it is unclear why political affiliation was differentially associated with speeded essentialist responding based on gender, liberal male participants did also exhibit an increased degree of change in responding between timing conditions. Generally we might have expected that while conservatives would exhibit greater essentialist responding overall, liberals may be pushed to be more essentialist when placed under time pressure due to having access to their political ideology restricted which would be consistent with these particular results. However, the finding that liberal males were in fact more essentialist makes this interpretation difficult.
Lastly, masculinity was positively associated with increased speeded essentialist responding for all participants and an increased degree of change in essentialist responding among male participants. Although no significant difference in males and females’ essentialist responding was observed generally, to the extent that masculinity can be conceptualized as capturing some culturally accepted attributes of “maleness,” this association is consistent with Eidson and Coley’s initial finding that males – though in this case those who identify as more masculine – exhibit greater essentialist responding. Thus, although we may not always see differences between male and female participants, gender role, as a conceptually similar identifier may be sensitive to differences in essentialism.

Notably, executive function measures were not associated with either essentialist responding or degree of change in responding for either male or female participants. That is, it does not appear that those who are better at inhibiting their gut responses simply exhibit fewer speeded essentialist responses nor are they simply more resistant to the time pressure manipulation. This suggests that a general inhibition explanation is not sufficient for explaining effects of time pressure on increased essentialist responding. However, this does not preclude more conceptual or social factors from triggering T2 processes which may suppress essentialist responding. Indeed, given that both cognitive measures (i.e., belief in variability) and social measures (i.e., motivation to avoid prejudice, political affiliation, and gender role) were predictive of essentialist responding or degree of change in responding in various cases but executive function measures were not, it seems plausible that Type 2 processes which inhibit essentialist responding are engaged through or are comprised of many separate factors.

Thus far, the Experiments 1 and 2 have shown that adults exhibit essentialist reasoning about gender categories that either increases when responses are given faster or the speed of
responding is experimentally manipulated, supporting the notion the dual-processing framework of essentialism. Further, the degree to which participants may be susceptible to this time pressure manipulation and the amount of essentialism they exhibit under it may be predicted by any number of individual difference measures in various instances. However, it is unclear if such a framework only applies to adults’ reasoning about gender categories or if such a framework can be applied to social categories more broadly. Thus, the next experiment will examine whether these findings are limited to reasoning about gender in particular or social categories generally while attempting to further discern the relationship between individual differences and essentialist responding.

**Experiment 3**

In experiment 2, Eidson and Coley’s (2014) initial finding that essentialist responding about gendered behaviors increases under time pressure was replicated within subjects demonstrating that time pressure can push individuals to respond in a more essentialist manner about gender. There were also some indications that these changes in essentialist responding may be predicted by other individual difference measures. Some of these may be more cognitive in nature such as beliefs in the variability of categorical properties, while some may be more social in nature such as varying levels of right-wing authoritarianism and subsequent reliance on or submission to authority.

Eidson and Coley, as well as the preceding two experiments presented here, have focused specifically on social essentialism and the category of gender. While the findings of these experiments suggest that social essentialism is merely inhibited in adults rather than replaced, it is unclear how true this may be of adults’ reasoning about social categories other than gender.
Thus, the aim of experiment 3 is to determine whether and to what extent social essentialism may persist for other types of social categories.

As has previously been reviewed, children are known to also maintain essentialist beliefs about racial categories (e.g., Hirschfeld, 1996; Rhodes & Gelman, 2009), ethnic and religious categories (e.g., Birnbaum et al., 2010; Deeb, 2011; Diesendruck & HaLevi, 2006), and gender categories (e.g., Rhodes & Gelman, 2009; Taylor, 1996; Taylor, Rhodes, & Gelman, 2009). In an objective sense, we may think of these social categories as ranging from more ontologically biological to ontologically social in nature. For example, the social category of gender is built upon the foundation of the biological category of sex. An individual’s sex, of course, is determined by the presence of either two X (female) or one X and one Y (male) chromosomes. While the social category of gender further heaps stereotypical behaviors upon this underlying biology, given gender’s identifiable biological roots it can be assumed that gender may objectively be viewed as a relatively more biological social category.

In contrast, social categories such as religious categories or political affiliations have no such discernible biological, chromosomal underpinnings – though religious categories may certainly be strongly linked with ethnicity in certain cases. Indeed, membership in such categories is objectively external in the sense that the categories are not necessarily inherent in the individual and are instead representative of external organizations. Further, an individual could presumably change their religious or political affiliation at any given time without changing anything about their own individual biological makeup. In this sense, we can view social categories such as religion or political affiliation as relatively less biological or more social in nature.
Given that social categories can range from more biological to more social in nature, an important question to be addressed in experiment 3 is whether or not essentialist responding occurs at the same base rate for a range of social categories. If, as proposed, essentialized social categories are treated as natural kinds regardless of their ontological nature (e.g., Atran, 1990; Gelman, 2003; Rothbart & Taylor, 1992), then to the extent that we observe essentialist responding at all for a given social category, it should occur at a relatively uniform rate. However, given that past work suggests that social categories can differ in the extent to which they are viewed as natural kinds (e.g., Haslam, Rothschild, & Ernst, 2000; Yzerbyt, 2001), we may observe variations in the levels of essentialist responding across a range of biological ontology for social categories. Specifically, Haslam et al found that adult participants rated both racial and gender categories as more natural kind-like than religious or political categories. Thus, on this view, we might expect to see greater essentialist responding overall for gender or racial categories compared to religious or political ones.

Similarly, should increases in essentialist responding occur under time pressure, it is unclear whether responding will increase at all or increase uniformly for all social categories that will be examined. It is possible that essentialist reasoning may be suppressed more for certain categories than others. If this were the case, we may see larger increases in essentialist responding for some categories and not for others.

Lastly, we must ask whether essentialist responding under time pressure can be predicted by individual difference measures. Although experiment 2 noted some relationship between individual difference measures and changes in essentialist responding about gender categories, it may be that different individual difference measures will be differentially predictive of the change in essentialist responding depending on the category being reasoned about. That is, do
different individual difference measures influence reasoning about categories in different ways? For instance, does one’s political affiliation affect reasoning about political affiliation categories? While such analogs were already in place for gender, race, and political affiliation, no such analog existed for religion in the existing set of individual difference measures. As such, a religiosity questionnaire was added based on the Fetzer Institute’s (1999) compilation of religiosity measures. In examining the possible mechanisms of Type 2 editing of essentialist Type 1 responses, any relationships between individual difference measures and changes in essentialist responding under time pressure may indicate a possible factor which influences reasoning outcomes in the essentialist dual processing system.

To answer these questions, we adapted the switched-at-birth task used by Eidson and Coley (2014) and others for use with four different social categories: race, gender, religion, and political affiliation. As in previous iterations of this experiment, participants reasoned about both behavioral and physical properties of social categories and did so under either time pressure or time delay. Switched-at-birth scenarios for race, religion, and political affiliation were identical in both the wording of the vignettes and the properties used except for the use of appropriate category labels. The gender switched-at-birth task was similar in content, however, as two same-sex parents alone cannot biologically produce a child and gender category membership would still be contingent on the gender of the child itself, the gender vignette described a single parent with same-sex children adopting an opposite-sex child (e.g., a single mother with several daughters adopting a baby boy). The gender scenario also used identical properties to the other category scenarios. Essentialist responding under different timing conditions and differences between them were assessed and subsequently analyzed alongside individual difference data in an attempt to better understand Type 2 editing of essentialist responses.
Method

Participants

One hundred and sixty Northeastern undergraduates completed the experiment for partial course credit for an undergraduate psychology course. Men and women were randomly assigned to reason about each of four social categories, resulting in approximately 40 participants reasoning about each category. Five participants’ data were excluded due to failure to follow procedural instructions. The resulting distribution of participants per category was as follows: race - 20 male, 18 female; gender - 19 male, 20 female; religion - 20 male, 22 female; political affiliation - 18 male, 18 female.

Design / Materials

Switched-at-Birth Task. While the switched at birth task was directly influenced by the gendered switched at birth task from other studies (e.g., Taylor 1996; Taylor Rhodes, & Gelman, 2009; Eidson & Coley, in press), it was adapted for use for all four social categories. Further, the wording of the vignette was changed to be slightly more realistic than past versions which were originally designed for use with children. A sample race vignette read as follows:

Two families (one who was white and one who was black) gave birth to infants in the same hospital on the same day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the white family brought home the child born to the black family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.
Please think about the child born to the black family but raised by the white family. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the black family *gave birth* to the child, but the white family *raised* the child.

For racial vignettes, participants randomly received two of six possible vignettes (one per timing condition) containing two of three possible racial categories (white, black, Asian) and were asked to reason about a child from either of the two families. Similarly, for religion vignettes, participants randomly received two of six possible vignettes (one per timing condition) containing two of three possible religious categories (Christian, Jewish, Muslim) and were asked to reason about a child from either of the two families. For political affiliation, participants randomly received one of two vignettes per timing condition reasoning about a child born to each of the two political affiliation categories (Democrat, Republican). Similarly, for gender categories, participants randomly received one of two vignettes per timing condition reasoning about an opposite-sex child adopted by either a single father and his several sons or by a single mother and her several daughters (see Appendix A for all possible vignettes).

Although the previous experiments examining the social category of gender used specific, stereotypical gendered properties, it was unclear, especially for behaviors, what an appropriately agreed upon stereotypical set of behaviors may be for the three new social categories. To address this issue, we decided to eschew specific instances of properties and instead use more general, abstract properties that attempted to capture the specific nature of properties used in Experiments 1 and 2. Because this led to participants being asked about a singular abstract property as
opposed to two specific stereotypical properties, participants were no longer asked to make a choice between two properties but instead to decide whether a child would share the property with its birth category or its adoptive category. Thus, after reading the vignette, participants randomly received 12 properties – six behaviors (e.g., “hobbies,” “preference for professions”) and six physical properties (e.g., “physical appearance,” “blood type”) and were asked to decide whether the child would share those properties with their birth category or adoptive category (see Table 4 for all properties). For example, a given trial might look like the following:

With which family do you think the child will share…

Hobbies

<table>
<thead>
<tr>
<th>White</th>
<th>Black</th>
</tr>
</thead>
</table>

Thus, participants would reason about one of the four social categories (between subjects) under both a time delay and speeded conditions (within subject). For each vignette, participants would reason about both behaviors and physical properties of a target (within subjects). This resulted in a 4 (category: race, religion, gender, political affiliation) by 2 (condition: delay, speeded) by 2 (property: behavior, physical) design.

**Individual difference measures.** Several of the same individual difference measures from experiment 2 – motivation to avoid prejudice, political affiliation, RWA, and executive function – were used with a few exceptions. First, the belief in category variability measure was not given to participants. This is due to the difference in property questions between the two experiments. Experiment 2 specifically asked participants to reason about stereotypical gendered properties, and, accordingly, the belief in category variability measure asked participants to rate the prevalence of those specific properties among both men and women. However, since no
specific stereotypical properties were used in this experiment, a category variability measure was not employed.

Religiosity. The Religious Meaning (Fetzer Institute, 1999) scale asked participants to rate how much meaning in their lives is derived from religious practice or affiliation (e.g. “My spiritual beliefs give meaning to life’s joys and sorrows,” or “My religious beliefs help me find a purpose in even the most painful and confusing events in my life”). While there are certainly many other ways to measure religiosity or religious belief, the Religious Meaning scale is notable for its lack of questions in reference to any particular belief system. Instead, it focuses on spiritual or religious beliefs broadly defined as they might relate to a participants’ life, making it useful for assessing religiosity in participants of any or no religious belief.

Procedure

Following the procedure laid out in previous experiments and by Eidson and Coley, participants completed the experiment in a laboratory setting using both SuperLab (Cedrus, 2006) and Qualtrics experimental presentation software (Qualtrics, 2013). First, participants completed either the Stroop or Semantic Inhibitory Control task. Following the first executive function measure, participants would randomly receive either a speeded or delayed switched at birth vignette. After reading the instructions, participant would randomly receive one of the possible combinations of vignettes appropriate for their given social category (e.g., for race, participants may reason about a child born to black parents but raised by white parents). Participants were given as much time as they wanted to read this vignette and were instructed to reason about one of the two children described in the story. Upon clicking next, participants began to randomly receive each of the twelve properties.
By pressing “f” participants chose the category that randomly appeared on the left side of the screen, while “j” would choose the category appearing on the right side of the screen. For participants in the delay condition, pressing a button before the 10 seconds had elapsed would not enter any response – only after the delay would the software make a record of participants’ choice. For participants in the speeded condition, if they did not respond within 2 seconds the screen would automatically move on to the next question, though participants would see the missed item again after all other items had been presented. If multiple items were not answered within the time limit, they would still be presented after all twelve original items were exhausted but in a new random order.

After completing the first switched at birth task, participants then completed all individual difference survey measures. Participants would receive one survey at a time in random order with that particular survey’s items also presented in a random order. All questions in the surveys were answered with Likert scales and participants merely clicked the appropriate response given a scale on the screen. After clicking next at the bottom of the screen, participants would then receive the next item. Once all individual difference surveys were completed, participants would receive the second timing condition with a new switched at birth vignette, and finally finish with the remaining executive function measure. This both temporally spaced the switched-at-birth tasks to avoid diminishing time pressure effects due to response familiarity and allowed participants time to recover from one executive function measure before completing the other.

**Results**

**Scoring**

**Switched At Birth.** For switched at birth tasks, participants were scored as having given an essentialist response if the category chosen matched the described child’s birth category. The
total number of essentialist responses was summed for both physical and behavioral properties in each of the timing conditions. Due to the nature of the experimental presentation software, some participants may not have answered all twelve property items in the speeded condition. To address this, the sum of a participant’s essentialist responses was divided by the total number of responses given per property type to attain the proportion of essentialist responses. Note that all of the following analyses were conducted with and without participants who answered all twelve items with no differences in results. As such, the results reported include all participants’ responses.

Individual differences. As in experiment 2, all individual difference measures were scored in accordance with their individual conventions. For survey measures, this involved averaging participants’ ratings to critical items to produce a single score for a given measure or construct (e.g., ratings for the motivation to avoid prejudice questions were averaged separately for internal and external motivation items to produce two separate average scores). For the executive function measures, this involved counting the number of errors made in critical conditions and subtracting errors in baseline conditions as well as recording the reaction times for correct responses.

Analyses

Essentialist Responses. Essentialist responding was conducted both for participants’ average essentialist responding and individual response patterns. We expected that essentialist responding about behaviors should show increases in both cases.

Mean Essentialist Responses. The switched at birth analyses address two of the main aims of experiment 3. Most importantly, we wanted to know if time pressure increased essentialist responding all social categories or just some. Secondly, we wanted to determine if
there were any differences in levels of essentialist responding among the four types of social
categories. To this end, we conducted a 4 (category: race, religion, gender, political affiliation)
by 2 (property: behavior, physical) by 2 (condition: delay, speeded) mixed measures ANOVA on
proportions of essentialist responding using category type as a between subjects variable and
property type and condition as within-subjects variables (see Figure 5).

Critically for our hypothesis, there was a property by condition interaction ($F(1,151) =
34.20, p < .001$). Planned paired-samples T-tests showed that significantly more essentialist
responses were given about behaviors under speeded conditions ($M = .221, SD = .252$) than
under delay conditions ($M = .107, SD = .173$), $t(154) = 6.526, p < .001$. Also, essentialist
responses about physical properties significantly decreased\(^2\) under speeded conditions ($M = .790,
SD = .212$) compared to delay conditions ($M = .846, SD = .156$), $t(154) = 2.961, p = .004$. For
both speeded and delay conditions, as was consistent with previous findings, significantly more
essentialist responses were given about physical properties than about behavioral properties,
Speeded: $t(154) = 19.025, p < .001$; Delay: $t(154) = 35.172, p < .001$. This increase in essentialist
responding about behavioral properties under time pressure provides additional evidence for the
dual-processing claim of essentialism.

Several tangential interactions and main effects were also observed. A main effect of
property ($F(1,151) = 1,118.146, p < .001$) showed that overall significantly more essentialist
responses were made about physical properties ($M = .818, SD = .144$) than behavioral properties
($M = .164, SD = .186$), while a main effect of condition showed that overall significantly more

---

\(^2\)Although the three-way interaction between category, condition, and property was not significant, it should be
noted that essentialist responses about physical properties only significantly decrease under time pressure compared
to time delay for race ($t(37) = 2.735, p = .010$) and religion ($t(41) = 2.068, p = .045$) but not for gender ($t(38) =
.972, p = .337$) or political affiliation ($t(35) = 1.700, p = .098$) suggesting that time pressure does not simply force all
response patterns toward chance.
essentialist responses were given under speeded conditions ($M = .506$, $SD = .140$) than under delay conditions ($M = .476$, $SD = .100$), and a main effect of category showed that overall significantly more essentialist responses were given about gender ($M = .557$, $SD = .142$) than about race ($M = .472$, $SD = .072$), religion ($M = .463$, $SD = .065$), or political affiliation ($M = .472$, $SD = .069$) – none of which significantly differed from each other.

A significant property by category interaction ($F(3,151) = 10.676, p < .001$) subsequently analyzed using two one-way ANOVAs showed an effect of category on overall behavioral responses ($F(3,151) = 15.963, p < .001$) with significantly more essentialist responses about behaviors regardless of timing condition for gender ($M = .316$, $SD = .231$) than race ($M = .133$, $SD = .157$), religion ($M = .127$, $SD = .143$), or political affiliation ($M = .075$, $SD = .091$)(all $ps < .001$) – though the latter categories did not significantly differ from each other (all $ps > .1$) – but no effect of category on essentialist responses about physical properties, $F(3,151) = 2.064, p = .107$. A significant condition by category interaction ($F(3,151) = 3.816, p = .011$) subsequently analyzed using two one-way ANOVAs showed an effect of category on overall speeded responses ($F(3,151) = 9.570, p < .001$) with significantly more essentialist responses under time pressure regardless of property for gender ($M = .601$, $SD = .184$) than race ($M = .466$, $SD = .113$), religion ($M = .472 SD = .092$), or political affiliation ($M = .483$, $SD = .110$)(all $ps < .001$) – though the latter categories did not significantly differ from each other (all $ps > .5$). There was also an effect of category on overall delayed responses ($F(3,151) = 2.804, p = .042$) with significantly more essentialist responses under a time delay for gender ($M = .513$, $SD = .142$) than religion ($M = .454$, $SD = .085, p = .008$) or political affiliation ($M = .461$, $SD = .076, p = .023$) but not more than race ($M = .478$, $SD = .072, p = .123$) – though race, religion, and political affiliation did not differ from each other, all $ps > .2$. 

ADULTS’ SOCIAL ESSENTIALISM
These effects generally suggest that essentialist responding measured by the SAB is greater overall for the category of gender compared to other categories. Although we might have predicted racial categories to have comparable levels of essentialist responding to gender categories, the finding that gender categories produce more essentialist responses than religious or political categories is consistent with the notion that social categories differ in degrees of perceived natural kind-ness.

**Individual Response Patterns.** Overall, the group results suggest that participants’ essentialist responding about social category-related behaviors increases under time pressure. Another way to examine this effect is through individual response patterns to determine if a significant number of individual participants changed their response patterns from delay to speeded conditions. As in Experiment 2, due to the repeated measures time pressure manipulation conventional Chi-Square analyses were not possible. Instead a McNemer’s Chi-Square test for paired data was used. This required that participants be typed into two nominal variables for two repeated measures. As an exploratory analysis, we set a cut-off of 60% essentialist responses for defining participants as having essentialist or non-essentialist response patterns. Thus, participants were typed as essentialists if they gave 60% or more essentialist responses about behaviors or non-essentialist if they gave less than 60% essentialist responses for both speeded and delay conditions. Counts were then submitted to a McNemer’s Chi-Square test to determine if significantly more participants exhibited consistent essentialist response patterns under speeded conditions \(N = 19\) than delay conditions \(N = 4\) (see Figure 6). Results showed that significantly different patterns of responses were obtained under speeded conditions than delayed conditions with more participants exhibiting consistent essentialist response patterns about behaviors under time pressure, \(p = .001\). This suggests that not only does
essentialist responding under time pressure increase in general, but that individuals are also more likely to give a set of consistent essentialist responses under time pressure compared to a delay.

**Individual differences.** As in Experiment 2, multiple regression analyses were conducted to determine if any relationship could be observed between various individual difference measures and essentialist responding – in this case, for both the change in essentialist responses from time delay to time pressure and speeded responses alone. Analyses were conducted separately for each social category and only for behavioral responses due to consistently high numbers of essentialist physical responses across categories and conditions. Regressions were performed on all participants’ degree of change and speeded behavioral responses but were also separated by male and female participants. The results of all regression analyses are summarized in Table 5 with individual predictors or models having $p < .09$ marked as at least marginally significant with significant effects for each predictor outlined below.

**Models.** Models were generally poor at predicting responding overall. However, taken together, the individual difference measures were able to predict change in essentialist responding about behaviors ($F(7,12) = 3.440, p = .029, R^2 = .667$) and speeded behavioral responses ($F(7,12) = 3.127, p = .040, R^2 = .646$) for male participants’ reasoning about race. When reasoning about religion, a significant model was found for all participants’ speeded behavioral responses ($F(7,33) = 2.399, p = .042, R^2 = .337$) and a marginally significant model for female participants’ speeded behavioral responses ($F(7,13) = 2.560, p = .068, R^2 = .580$)

**Motivation to Avoid Prejudice (Internal / External).** Internal motivation to avoid prejudice was found to be a predictor in two instances – for female participants, internal motivation to avoid prejudice was marginally associated with increased essentialist responding when reasoning about gendered behaviors ($\beta = .571, p = .082$) and significantly associated with
an increased degree of change in essentialist responding about when reasoning about religious behaviors (\( \beta = .557, p = .042 \)). External motivation to avoid prejudice was not found to be a significant predictor of either degree of change in essentialist responding or speeded behavioral responses for any category or subset of participants.

**Political Affiliation.** Political affiliation was found to be a significantly associated with participants’ degree of change in essentialist responding about racial behaviors (\( \beta = -.622, p = .006 \)) with liberal participants showing greater degrees of change. Political affiliation was also marginally predictive of female participants’ degree of change in essentialist responding about behaviors for race (\( \beta = -1.108, p = .057 \)) – but not for male participants (\( \beta = -.377, p = .156 \)) – with more liberal women showing greater degrees of change.

**RWA.** RWA was predictive in only one instance. RWA was positively associated with participants’ degree of change in essentialist responding about racial behaviors (\( \beta = .562, p = .030 \)).

**Gender Role (Masculinity / Femininity).** Masculinity was marginally predictive of the degree of change in essentialist responding about behaviors (\( \beta = .411, p = .082 \)) and significantly predictive of speeded behavioral responses (\( \beta = .493, p = .044 \)) for male participants’ reasoning about race. Masculinity was also marginally predictive of participants’ speeded behavioral responses when reasoning about gender (\( \beta = .297, p = .075 \)). Alternatively, femininity was significantly negatively associated with male participants’ degree of change in essentialist responding about behaviors when reasoning about race (\( \beta = -.583, p = .022 \)) and both all participants’ (\( \beta = -.436, p = .019 \)) and female participants’ (\( \beta = -.559, p = .018 \)) speeded behavioral responses when reasoning about religion.
Religiosity. Religiosity significantly predicted degree of change in essentialist responding about behaviors when reasoning about race ($\beta = .650, p = .048$) and marginally predicted speeded behavioral responses when reasoning about religion ($\beta = .537, p = .090$) only for male participants.

Discussion

As previously noted, there were three specific aims of experiment 3. First, we wanted to know whether or not Eidson and Coley’s and previous experiments’ finding that essentialist responding about gendered behaviors increased under time pressure would generalize to other social categories. Second, given that some social categories can be objectively identified as ontologically more biological or perceived as more natural-kind like than others, are there differences in essentialist responding along such a spectrum? And third, we wanted to know if individual difference measures could predict the degree of change of essentialist responding or number of speeded essentialist responses in order to better understand elements of the Type 2 editing process which seemingly inhibits essentialist Type 1 responding within the proposed essentialist dual processing framework.

The results of experiment 3 showed that Eidson and Coley’s initial finding that essentialist responding about gendered behaviors increases under time pressure does indeed generalize to other social categories. In fact, for all of the social categories examined in experiment 3 (race, religion, gender, political affiliation), essentialist responding about behaviors significantly increased under time pressure. This occurred even though experiment 3 employed a more realistic switched-at-birth vignette than had been used in previous studies and that the properties participants were asked to reason about were abstract in nature as opposed to specific
stereotypical properties associated with a given social category. Still, despite this, we find a strong repetition of the time pressure effect across all four social categories.

This increase in essentialist responding under time pressure supports the claim that the essentialist bias is both fast and automatic bias and persists into adulthood rather than changing over the course of development. By placing our adult participants under time pressure we see an increase in essentialist responding comparable to that of 10-year old children, at least for gender categories (Taylor, Rhodes, & Gelman, 2009). This is true not only for the social category of gender, but also for a range of social categories – though the increase in responding for other categories did not reach the same levels as gender. This may suggest that when unimpeded the essentialist bias can at begin to affect reasoning about a wide range of social groups.

The results did show, however, that essentialist responding about gender appears to occur at higher rates under both delay and speeded conditions than for other social categories. As previously noted, gender is certainly a social category that may be both objectively more biological and natural-kind like than other social categories like religion or political affiliation but it also perceived that way (Haslam et al, 2000). In this sense, it is perhaps not surprising that we would see increased base rates of essentialist responding. However, we do not see similar levels of responding for race – a social category which shares both perceived and objective natural-kindness with gender. Thus, while this is certainly in line with Yzerbyt (2001) and Haslam and colleague’s (2000) proposal that certain social categories may be more essentialized than others depending on varying perceptions of natural kindness and entitativity, it may also point to gender being a privileged social category within our cultural milieu with respect to its inherent potential.
Lastly, we noted that at least some individual difference measures were able to predict both changes in essentialist responding between different timing conditions and speeded responding alone. Surprisingly, internalized motivation to avoid prejudice (i.e., an individual personally believes that stereotyping should be avoided) was positively associated with essentialist responding for female participants’ reasoning about gendered behaviors. This is contrary to Plant and Devine’s (1998) past findings suggesting that motivation to avoid prejudice—especially internal motivation—is associated with decreased stereotyping and presumably by analogy essentialist responding. However, internal motivation to avoid prejudice was also positively associated with the degree of change between delayed and speeded responses when female participants were reasoning about religious behaviors. To the extent that degree of change can be conceptualized as active suppression of essentialist responding, this relationship is consistent with the notion that those with higher internal motivation to avoid prejudice give fewer essentialist responses under time delay when they have time to edit their response and more essentialist responses under time pressure when they do not. This provides mixed evidence as to the general effect that pragmatic social factors may have on essentialist responding, but suggests that in certain instances social motivation certainly can lead to editing of essentialism though it may not always.

Like motivation to avoid prejudice, political affiliation showed a significant relationship with the degree of change in essentialist responding about racial behaviors between delayed and speeded responses. Specifically, participants who considered themselves to be more liberal (which has been shown to be associated with decreased stereotyping—e.g., Napier & Jost, 2008; Sidanius, Pratto, & Bobo, 1996) were more likely to show a greater degree of change in essentialist responding. Conversely, participants with higher RWA (which has been shown to be
associated with increased stereotyping – e.g., Altemeyer 1981; 1996, 2006) were more likely to show a greater degree of change in essentialist responding. Thus it is difficult to say how these two constructs related to political ideology may generally be related to the suppression of essentialist responding from these data alone.

Analogous to Eidson and Coley’s (2014) finding that overall male participants give more essentialist responses than female participants, masculinity (to the degree that it can be considered to capture some aspect of stereotypical maleness) was associated with increased essentialist responding for all participants’ reasoning about gendered behaviors and for male participants’ reasoning about racial behaviors. It was also associated with an increased suppression of essentialist responding with masculinity being associated with a greater degree of change in essentialist responding between timing conditions. Similarly, for both participants generally and female participants specifically femininity was negatively associated with degree of change in essentialist responding about religious behaviors and negatively associated with degree of change in essentialist responding about racial behaviors. This may suggest that gender roles which align more with maleness (i.e., masculinity) or away from femaleness (i.e., femininity) may exhibit greater essentialist responding analogous to general gender effects that have been previously reported.

Finally, religiosity was shown to be marginally associated with increased essentialist responding for male participants’ reasoning about religious behaviors and their degree of change in essentialist responding when reasoning about racial behaviors. Kelemen (2003) noted that increased cultural religiosity was associated with more promiscuous teleological responding in American children compared to British children – though British children still exhibited overall high amounts of teleological endorsement. As teleology is also an early-emerging conceptual
bias that seems to persist into adulthood and can be increased under time pressure, it is entirely consistent that similar effects of religiosity would be seen on essentialist responding with increased religiosity resulting in increased essentialist responding.

It should be noted, however, that overall no one individual difference measure was predictive of essentialist responding. Further, those measures that were predictive were rarely associated with essentialist responding for more than one social category or for both male and female participants. This provides further evidence that T2 editing of essentialist responses may be influenced by multiple factors based on reasoning context – whether that may be the category being reasoned about, the gender of a participants, or any other factor.

Overall, experiment 3 showed that essentialist responding does appear to occur and be increased by time pressure for a range of social categories. Further, although there are some notable differences in base rates of essentialist responding (specifically for the category of gender), the time pressure effect noted by Eidson and Coley (2014) uniformly applies to both more and less biological social categories providing evidence for the dual-processing account of essentialism. There is also indication that changes in essentialist responding and speeded responses may be partially predictable for some categories by means of different individual difference measures. Subsequent experiments examined how generalizeable these findings are beyond the scope of the switched-at-birth task by examining essentialist responding in other experimental contexts.

**Experiment 4**

The previous three experiments have shown that while essentialist responding about physical properties is ubiquitous, essentialist responding about behavioral properties of social categories, while low among young adults at baseline, increases under time pressure. Those
experiments have supported a dual processing account of social essentialism. However, all of these results have been obtained using a SAB task which taps into only one aspect of essentialist thinking. While Experiments 1-3 focused on the inherent potential of social categories and a dual processing framework of essentialism, this experiment and Experiment 5 will explore and assess the greater extent of adults’ social essentialism. As such, experiment 4 is an initial investigation of another aspect of essentialist reasoning using a social transformation task.

In contrast to the inherent potential of categories assessed by the first three experiments in this dissertation, beliefs about resistance to change tap a different aspect of essentialist reasoning. Inherent potential is an underlying belief that regardless of a young child’s appearance or traits, the underlying category essence will give rise to specific category-based properties over the course of development. Underlying beliefs about resistance to change are instead a function of the assumed causal nature of essences. That is, if essences define category membership, are stable over time, and are the underlying primary cause of category-based properties, changes to less causally central properties like appearance will not greatly impact beliefs about category membership. Thus, regardless of various changes a category member may be subject to, category membership is maintained do to the continued presence of an unaltered essence.

As previously outlined in the general introduction, transformation tasks have focused on the ways in which adults and children understand the stability of natural kind category membership (e.g., Frisson & Wakefield, 2012; Gelman & Wellman, 1991; Hampton, Estes, & Simmons, 2007; Keil, 1989; Rips, 1989). In these studies, participants are often told that some type of change has occurred for a particular animal. These may be purely appearance-based changes such as a horse being painted to look like a zebra (e.g., Frisson & Wakefield, 2012) or internal changes such as scientists replacing the “insides” of one animal with those of another
(e.g., Gelman & Wellman, 1991). The former generally results in participants preserving
category membership and assuming that a coat of paint is not sufficient to change or remove the
underlying essence, while the latter often results in participants reporting that category
membership has indeed changed. These studies suggest that at least for reasoning about animal
kinds, properties like appearance have comparatively weaker causal links to a category essence
than do internal properties. Thus, when critical properties with strong causal links to the essence
– in this case internal properties – are transformed, the essence may be assumed to have been
removed and replaced leading participants to believe a change in category membership has
occurred.

While this work suggests that transformations of different properties may differentially
affect beliefs about category membership, there has been little investigation as to how
individuals may perceive social transformations. One notable exception involves work conducted
by Mahalingam (2003) in which she examined beliefs about gender transformations in India
among the Aravanis – a community of mostly female transgendered individuals who generally
wear female clothing and often undergo sex-change operations to change from being
anatomically male to anatomically female. In general, the Aravanis endorsed beliefs that one
could change from male to female but not necessarily change from female to male. Further, both
superficial changes such as wearing female clothing and surgical operations such as a sex change
were seen as able to perpetuate such a change. Mahalingam’s work suggests that some social
categories may be viewed as stable and unchanging – like the category female – but there may
also be membership in some categories which is seen as easily malleable – like the category male
– at least among a very specific subset of the population.
Other evidence suggest that social categories may vary in their malleability. Haslam, Rothschild, and Ernst (2000)’s essentialist dimension of naturalness encompasses concepts of immutability (i.e., the degree to which category membership could be easily changed) and stability (i.e., the degree to which category membership was stable over time) among others. Categories which are rated highly on the naturalness dimension should then be less susceptible to transformations given their perceived stability and immutability. In Haslam’s study, categories like race and gender tended to receive stronger naturalness endorsements while categories like religion and political affiliation received much weaker endorsement. Thus, we might expected that race and gender should be less susceptible to transformations overall compared to religion and political affiliation.

Aside from overall susceptibility to transformation, Haslam et al’s dimensions of essentialism may also predict differential effects for different types of transformations. Given that naturalness refers to the extent to which categories can be viewed as naturally occurring or are natural kind-like while cohesiveness refers to the extent to which categories and their members share similar goals and behaviors, we might predict that transformations to physical or biological properties (similar to the “insides” used by Gelman & Wellman, 1991) might affect categories perceived as more natural while transformations to characteristic behaviors (similar to wearing dresses or exhibiting stereotypical female behaviors used by Mahalingam, 2003) might affect more cohesive categories. Thus, given their comparatively high naturalness and low cohesiveness, racial and gender category membership may be more susceptible to transformations of physical properties while religious and political categories which were rated highly on cohesiveness but much lower on naturalness may have membership be more susceptible to transformations of characteristic behaviors.
Experiment 4 will thus examine participants’ views about gender, racial, religious, and political transformations to determine whether these categories have critical properties that once changed produce greater beliefs about moving from one category to another. If so, this may suggest that such properties are more causally central to category membership and thus indicative of essentialist reasoning. These transformations will include stereotypical behavioral changes and physical changes (e.g., surgical alteration of genitalia), or both for natural categories (i.e., race and gender) to examine what types of changes, if any, can cause a shift in social category membership. However, as more cohesive categories (i.e., religion and political affiliation) lack such salient physical features, transformations to these categories were presented as occurring over outward characteristic behaviors (e.g., going to church or attending a political rally) and nominal category properties – properties that allowed objective identification with a particular category (e.g., one’s voter registration). As previously noted, these four categories vary in their perceived naturalness and as such may differ in the extent to which membership in these categories is affected by changes to any property overall, while each category’s respective degrees of naturalness and cohesiveness could predict differential effects for each type of property transformation.

Although no time pressure manipulation was employed, the dual processing account of essentialism was examined, as in experiment 1, by examining the response times in various conditions. Given Haslam et al’s proposal that racial and gender categories more align with the naturalness aspect of essentialism, we might expect that not only that they may be affected by physical transformations but that reasoning about them in the physical transformation condition should take longer than for stereotypical behavioral transformations. Conversely, for religious and political categories which are rated more highly on cohesiveness and thus may be more
susceptible to characteristic behavioral transformations, we might expect that responses in the characteristic behavior condition should take longer than for nominal transformations. Such timing evidence would suggest that responses given under more amenable essentialist conditions are the result of a fast, automatic T1 bias when they are produced without interference from T2 process.

Further, the individual difference measures employed in experiments 2 and 3 will be used to examine whether willingness to accept social transformations can be predicted by other factors (refer back to Table 2 for predicted directions of associations).

Method

Participants

128 participants (64 female) were recruited from Northeastern University’s undergraduate psychology pool. Participants were randomly assigned to one of three transformation conditions: 43 participants (22 female) reasoned about changes to stereotypical and characteristic behaviors, 43 participants (21 female) reasoned about physical and nominal transformations, and 42 participants (21 female) completed the “both” condition.

Materials and Design

Norming. Properties intended for use in the transformation task were first assessed for their perceived underlying cultural stereotypicality in order to determine whether changes to more stereotypical properties simply resulted in a greater change in category membership regardless of the property being reasoned about. Ruling out this possibility ensures that what is critical is the type of property being transformed rather than simply differences in cultural expectations. A separate group of 18 Northeastern University undergraduates were told that cultural stereotypicality referred to the extent that a property was seen as part of a cultural belief
about of members of a given category even if they personally felt it may not be true or that they
did not actually know a member of a category who exhibited said property (e.g., “Professors are
absent-minded.”). They were then asked to rate the cultural stereotypicality of properties
intended for use for a particular social category on a scale from 1 to 9 with higher ratings
representing greater stereotypicality (e.g., for gendered properties participants rated how
culturally stereotypically male and female properties were of males and females separately).

Paired-samples t-tests showed that in all instances, properties were rated as more
stereotypical of their intended category than of the opposing category (e.g., stereotypical black
properties were rated as significantly more stereotypical of black people than of white people)
(all ps < .001). Thus for individual categories, ratings for intended categories were considered
together (e.g., intended stereotypical black behaviors and intended stereotypical white behaviors
were subsumed under stereotypical racial behaviors). Paired-samples t-tests were then used to
assess whether properties were considered equally stereotypical for that category alone. Results
showed no difference in the stereotypicality of racial behaviors (M = 7.148, SD = 1.638) or
physical properties (M = 7.167, SD = 1.750) (p = .913) or in the stereotypicality of characteristic
behaviors (M = 6.968 SD = 1.902) or nominal properties of religion (M = 7.069, SD = 2.168) (p
= .664). However, physical properties of gender (M = 8.056, SD = 2.301) were rated as more
stereotypical than behaviors (M = 7.213 SD = 1.764) (p < .001) and nominal properties of
political groups (M = 7.528, SD = 1.824) were rated as more stereotypical than characteristic
behaviors (M = 6.148, SD = 1.725) (p < .001). Thus, the critical test of whether changing more
stereotypical properties results in greater change in category membership can be observed from
beliefs about change for gender and political groups.
**Transformation Task.** Modeled after Rips’s (1989) and Hampton and colleagues’ (2007) adult transformation task, participants first read a short description of an individual who had undergone one of three changes depending on the category: changes to stereotypical or characteristic behaviors, changes to physical or nominal properties, or both (see Table 6 for a summary of items for each condition). Participants reasoned about changes between racial (black/white), gender (male/female), religious (Christian/Muslim), and political affiliation (Democrat/Republican) categories in random order (see Appendix B for all vignettes). Participants were asked to determine the categorical identity of the individual both pre- and post-transformation. For example, participants may have read the following short vignette:

Mike was a healthy young man. He liked to wear shirts and ties, watch football, and drink beer. One day he decided to change to be like a young woman instead. He had his genitals surgically transformed into those of a young woman and started taking hormone treatments. Mike continued to wear shirts and ties, watch football, and drink beer.

(Before / after) Mike changed would you consider Mike more of a male or more of a female?

1(definitely male)  2  3  4  5  6  7  8  9(definitely female)

Thus, participants received one of three types of transformations between-subjects, and within-subjects participants reasoned about each of four social categories (gender, race, religion, political affiliation).
**Individual Differences.** Due to the relative lack of a predictive relationship between general executive function measures and essentialist responding, the Stroop task and SIC were dropped from the battery for brevity. Otherwise, participants completed the same battery of individual difference measures used in Experiment 3 – motivation to avoid prejudice, political affiliation, RWA, gender role as measured by the BSRI, and religiosity.

**Procedure**

Participants completed both the transformation task and individual difference measures on a computer using SuperLab (Cedrus, 2006) and Qualtrics survey software (Qualtrics, 2013), respectively. After being given instructions on the structure of the transformation task, participants received an appropriate vignette based on condition for each social category in random order. For each vignette participants were asked to assess the individual’s category identity both pre- and post-transformation with the order of category judgments randomized for each social category. Participants proceeded through each vignette and subsequent responses at their own pace. After receiving all transformation vignettes participants received each of the individual difference measures in random order and received each of the items for those measures in random order.

**Results**

**Scoring**

For each category, participants made both a pre- and post-transformation category judgment indicating which social category they felt that each individual described in the vignettes belonged to. In this case, essentialist responding is dependent on both the pre- and post-transformation but not necessarily on the absolute category ratings. Participants could indicate that Mike was a male both before and after his transformation. This preservation of category
membership is an essentialist response. However, participants could indicate based on the vignette that Mike had actually been a female all along before his transformation and subsequently was still female. For the purposes of this experiment, this is also an equivalent essentialist response.

Thus, a difference score was taken to indicate the extent to which transformations affected category membership. For instance, if a participant rated the individual in the example above as being definitely female (a response of 9) both pre- and post-transformation, the difference would be zero – indicating no effect of the transformation scenario on category judgment. However, if the participants rated the individual as exceptionally male pre-transformation (a response of 1) but exceptionally female post-transformation (a response of 9), the difference from pre- to post-transformation would be eight, indicative of a strong transformation effect.

Reaction times were also recorded for pre- and post-transformation category judgments and difference scores were computed in the same manner as the category judgments. Individual difference measures were scored according to their conventions outlined in Experiment 2.

Analyses

Category Judgments. The main question of this experiment was whether social transformations cause participants to change how they categorize members of different social categories. Further, we were interested in whether certain social categories may be more susceptible to different types of social transformations. We generally hypothesized that categories which are perceived to be more natural like gender and race would be less susceptible to transformations overall but also that they would be more affected by physical transformations than changes to stereotypical behaviors. Conversely, categories which are perceived to be more
cohesive like religion and political affiliation would be more affected by changes to characteristic properties than nominal ones.

**General Susceptibility to Transformation.** To first test whether social transformations affected category judgments for social categories at all, differences in category judgments for each category were subjected to a one-sample t-test against zero to determine if social transformations, in general, affect category membership. Each category – race ($M = 2.156, SD = 2.312$), gender ($M = 4.164, SD = 2.525$), religion ($M = 3.859, SD = 3.040$), and political affiliation ($M = 3.906, SD = 2.947$) – all significantly differed from zero regardless of transformation condition (all $p < .001$). This suggests that all of the categories examined are at least somewhat affected by transformations overall.

**Differential Transformation Effects.** Next, we wanted to determine if categories perceived to be more natural would be less susceptible to transformations overall than less natural ones and to test whether different social categories were differentially susceptible to different types of transformations. To answer these questions we ran a 3 (transformation condition: stereotypical / characteristic behaviors, physical / nominal properties, both) by 4 (category: race, gender, religion, political affiliation) mixed measures ANOVA with transformation condition as a between-subjects variable, category as a within-subjects variable, and differences in category judgments as the dependent variable.

Critically for our hypotheses, the ANOVA showed a significant interaction between category and transformation condition suggesting that indeed ratings of category membership are differentially affected by certain transformations for different categories, $F(6,375) = 7.257, p < .001$, (see Figures 7 and 8). In order to tease apart the interaction, separate repeated measures ANOVAs were run for each transformation condition using category as the independent variable
and the differences in category judgments as the dependent variable. Separate one-way
ANOVAs were also run for each category using transformation condition as the independent
variable and the differences in category judgments as the dependent variable. Both
interpretations of the interaction are reported below – though it should be noted that comparisons
within transformation condition may be greatly influenced by the different types of properties
used for each category.

For the stereotypical / characteristic behavioral transformation condition, we predicted
that more cohesive categories – religion and political affiliation – would show greater
susceptibility to transformation while less cohesive categories – race and gender – would show
less susceptibility to transformation. The ANOVA revealed an effect of category \((F(3,126) = 22.997, p < .001)\) with Tukey LSD pairwise comparisons showing that participants viewed race
\((M = 1.395, SD = 1.954)\) as significantly less susceptible to change than gender \((M = 3.465, SD = 2.491)\), religion \((M = 4.372, SD = 3.338)\), and political affiliation \((M = 4.442, SD = 3.142)\) (all \(ps < .001)\). Gender was also viewed as less susceptible to change than religion and political
affiliation \((ps = .023 and .019, respectively)\), though religion and political affiliation did not
differ in their susceptibility to transformation \((p > .8)\). Thus, the stereotypical / characteristic
behavioral indeed affected category judgments about more cohesive categories the most while
less cohesive categories showed comparatively less change in category membership – though
race exhibited even less change than gender.

For the physical / nominal property transformation condition, we expected that more
natural categories – race and gender – would be the most susceptible to changes in category
membership while less natural categories – religion and political affiliation – would be the least
susceptible to transformation. The ANOVA revealed an effect of category \((F(3,126) = 13.295, p\)
< .001) with Tukey LSD pairwise comparisons showing that participants viewed race ($M = 2.442$, $SD = 2.302$), religion ($M = 2.605$, $SD = 1.761$), and political affiliation ($M = 1.698$, $SD = 2.099$) as significantly less susceptible to change than gender ($M = 4.512$, $SD = 2.240$) (all $ps < .001$). However, race, religion and political affiliation did not differ in their susceptibility to transformation (all $ps > .55$). This indicated that gender – a more natural category – was the most susceptible to its physical transformation while less natural categories like religion and political affiliation were less affected by their nominal transformation. However, although race is perceived to be a more natural category, it was comparatively less affected by transformation.

For the “both” transformation condition, there was an effect of category ($F(3,123) = 7.474$, $p < .001$) with Tukey LSD pairwise comparisons showing that participants viewed race ($M = 2.643$, $SD = 2.507$) as significantly less susceptible to change than gender ($M = 4.524$, $SD = 2.734$), religion ($M = 4.619$, $SD = 3.393$), and political affiliation ($M = 4.595$, $SD = 3.155$) (all $ps < .001$). However, gender, religion, and political affiliation did not differ in their susceptibility to the “both” transformation (all $ps > .85$). This indicates that overall race was still more resistant to change than other social categories.

Conversely, breaking the interaction down by category revealed that for race, there was a significant effect of condition ($F(2,125) = 3.742$, $p = .026$) with pairwise comparisons showing that participants judged there to be significantly less change in the stereotypical behavioral condition than the physical condition ($p = .034$) and the both condition ($p = .012$), though there was no significant difference between the physical and both conditions. Similarly, there was a marginally significant effect of condition for gender ($F(2,125) = 2.542$, $p = .083$) such that participants judged there to be marginally significantly less change in the stereotypical behavioral condition than the physical condition ($p = .054$) and the both condition ($p = .053$),
while the physical and both condition did not significantly differ from each other ($p = .982$). This confirms our prediction that categories which are perceived to be more natural are more affected by physical transformations. Further, there is no additive effect from combining the physical transformation with the stereotypical behavioral transformation.

In contrast, an opposite pattern was observed for less natural and more cohesive categories. There was a significant effect of condition for religion ($F(2,125) = 6.029, p = .003$), participants judged there to be significantly less change in the nominal property condition than the characteristic behavioral condition ($p = .006$) and the both condition ($p = .002$), though the characteristic behavioral and both condition did not differ from each other ($p = .698$). Judgments for political affiliation followed a similar pattern with a significant effect of condition ($F(2,125) = 5.897, p = .004$) such that participants judged there to be significantly less change in the nominal property condition than the characteristic behavioral ($p = .005$) and the both condition ($p = .003$), though the characteristic behavioral and both condition did not differ from each other ($p = .804$). This suggests that membership in more cohesive categories is more subject to characteristic behavioral transformations and that there is no additive effect from combining the characteristic behavioral with the nominal transformation.

The ANOVA also revealed a main effect of category ($F(3,375) = 26.94, p < .001$) with pairwise comparisons showing that participants were less willing to accept social transformations for race ($M = 2.156, SD = 2.312$) than all other social categories (gender: $M = 4.164, SD = 2.525$; religion: $M = 3.859, SD = 3.040$; political affiliation: $M = 3.906, SD = 2.947$; all $ps < .001$), though gender, religion, and political affiliation did not significantly differ from each other (all $ps > .2$). There was also a marginal main effect of transformation type ($F(2,125) = 2.928, p = .057$), with pairwise comparisons showing that participants were more willing to accept social
transformations under the “both” condition ($M = 4.070$, $SD = 3.037$) than the physical / nominal condition ($M = 3.065$, $SD = 2.270$) ($p = .019$), though the “both” condition did not significantly differ from the stereotypical / characteristic behavioral condition nor did the stereotypical / characteristic behavioral condition significantly differ from the physical / nominal condition ($ps = .120$ and .411, respectively). However, these effects are subsumed under the interaction outlined above.

Taken together, these findings suggest that while overall race is slightly less susceptible to transformation than other social categories, different categories are susceptible to different types of changes. More natural categories like race and gender are more affected by changes to physical properties than stereotypical behaviors, while more cohesive categories like religion and political affiliation are more susceptible to changes in characteristic behaviors than nominal properties. Further, there is no observed additive effect from the two types of change with the “both” condition yielding similar rates of change as the physical condition for race and gender and the characteristic behavioral condition for religion and political affiliation.

**Individual Response Patterns.** In addition to group-level analyses, examining individual response patterns can provide insight into the effects of each transformation condition on individual participants. Unlike individual analyses in Experiments 2 and 3, participants were typed into three separate groups – essentialist, some change, and non-essentialist – characterized by category difference scores from zero to two, three to five, and six to eight, respectively. Analogous to group-level analyses, we would expect that when reasoning about race or gender, there should be more participants exhibiting essentialist response patterns under stereotypical behavioral transformation conditions than under physical or both conditions and more “some change” or non-essentialist response patterns under physical and both conditions than under
stereotypical behavioral conditions. Alternatively, when reasoning about religion or political affiliation, we would predict more participants to exhibit essentialist response patterns under nominal property conditions compared to characteristic behavioral or both conditions and more “some change” or non-essentialist response patterns under characteristic behavioral and both conditions compared to the nominal property condition.

Counts of response patterns for each transformation condition were subjected to chi-square analyses for each category separately (see Figures 9-12). Chi-square analyses were not significant for race ($\chi^2 = 6.580, p = .160$) or gender ($\chi^2 = 7.030, p = .134$). However, counts of responses did pattern in the expected direction with more participants exhibiting essentialist response patterns under stereotypical behavioral transformation conditions and more participants exhibiting non-essentialist response patterns under physical and both transformation conditions. Analyses did reveal significant effects for religion ($\chi^2 = 30.471, p < .001$) and political affiliation ($\chi^2 = 28.474, p < .001$), however, with more participants exhibiting essentialist response patterns under nominal property conditions than the characteristic behavioral or both conditions and more participants exhibiting non-essentialist response patterns under characteristic behavioral and both conditions than the nominal condition for both categories. This is consistent with the group-level analyses showing that categories like race and gender which are perceived to be natural but not cohesive are most susceptible to physical changes and least susceptible to stereotypical behavioral changes and that categories like religion and political
affiliation which are perceived to be cohesive but not natural are more affected by characteristic behavioral changes but not by nominal changes.

**Timing effects.** Although this experiment was designed to primarily investigate other aspects of adults’ social essentialism rather than directly testing the dual-processing account outlined in experiments 1 through 3, reaction times were collected for each category judgment in order to determine whether participants took different amounts of time to respond based on transformation condition. We predicted that response times would be longer for race and gender categories in the physical or both condition where changes in category rating were most likely to occur compared to the stereotypical behavioral condition. Alternatively, we predicted that response times would be longer for religion and political affiliation in the characteristic behavioral or “both” condition where changes in category rating were likely to occur compared to the nominal conditions.

Differences in pre-and post-transformation reaction times were analyzed in a 3 (transformation condition: stereotypical / characteristic behaviors, physical / nominal properties, both) by 4 (category: race, gender, religion, political affiliation) mixed measures ANOVA with condition as a between-subjects variable and category as a within-subjects variable. However, no significant effects or interactions were observed (all $ps > .150$). These results may suggest that the transformation task only relies on one type of reasoning – either Type 1 or Type 2 – to produce responses, thus failing to produce observable reaction time differences.

**Individual differences.** As in previous experiments, multiple regression analyses were conducted to determine if any relationship could be observed between various individual difference measures and essentialist responding – in this case, differences in pre- and post-
transformation category judgments. Analyses were conducted separately for each social category. Regressions were performed on all participants’ judgments about a given category’s transformation ratings and male and female’s transformations separately. Regressions were also performed for each transformation condition separately for the appropriate subset of participants and then for male and female participants in each transformation condition separately. For the purposes of interpreting associations with individual differences, recall that because difference scores were used, differences closer to zero represent a belief in less susceptibility to transformation – an essentialist response – while greater difference scores represent a belief in greater susceptibility to transformation – a non-essentialist response. Thus, positive associations indicate that the measure is associated with decreased essentialist responding while negative associations indicate that the measure is associated with increased essentialist responding. The results of all regression analyses are summarized in Tables 7 and 8 with individual predictors or models having $p < .09$ marked as at least marginally significant.

Models. As in previous experiments, models were generally poor at predicting essentialist responding overall. However, taken together, the individual difference measures were able to predict differences in male participants’ categorical judgments about race in the stereotypical behavioral transformation condition ($F(7,20) = 2.931, p = .047, R^2 = .612$), participants’ overall differences in categorical judgments about gender ($F(7,127) = 2.185, p = .040, R^2 = .113$), and participants’ differences in categorical judgments about gender in the “both” transformation condition ($F(7,34) = 2.643, p = .027, R^2 = .352$) and marginally predict participants’ differences in categorical judgments about gender in the physical transformation condition ($F(7,35) = 1.955, p = .027, R^2 = .137$).
**Motivation to Avoid Prejudice (Internal / External).** Internal motivation to avoid prejudice marginally predicted male participants’ reasoning about racial transformations, $\beta = -.649$, $p = .060$. This indicated that males with self-reported as having internal pressure to avoid stereotyping produced more essentialist responses counter to expectations that motivation to avoid prejudice should be associated with decreased essentialist reasoning.

However, external motivation to avoid prejudice also significantly predicted male participants’ reasoning about racial transformation ($\beta = .586$, $p = .042$) and marginally predicted female participants’ reasoning about nominal political transformations ($\beta = .407$, $p = .079$). In both cases, external motivation to avoid prejudice was associated with decreased essentialist responding. This is consistent with the prediction that motivation to avoid prejudice should result in less essentialist responding.

**Political Affiliation.** Political affiliation was a significant predictor only for participants’ reasoning about gender transformations in the physical condition, $\beta = .423$, $p = .048$. Specifically, conservatism was associated with decreased essentialist responding (or a greater difference in category ratings). This was counter to the prediction that conservatism would be associated with increased essentialist responding.

**RWA.** RWA was only marginally predictive of male participants’ reasoning about racial transformations in the stereotypical behavioral condition, $\beta = -.671$, $p = .071$. Specifically, RWA was associated with increased essentialist responding (or a smaller difference in category ratings). This was consistent with the prediction that RWA would be associated with increased essentialist reasoning.

**Gender Role (Masculinity / Femininity).** While masculinity was not predictive of participants’ reasoning for any category or transformation condition, femininity was predictive in
several instances. Femininity was a significant predictor of participants’ reasoning about racial transformations in the stereotypical behavioral condition ($\beta = -.392, p = .040$), male participants’ reasoning about racial transformations in the stereotypical behavioral condition ($\beta = -.696, p = .005$), and marginally predictive of female participants’ reasoning about political transformations in the nominal condition, $\beta = .533, p = .051$. In the former two cases, femininity was associated with increased essentialist responding, but in the latter instance femininity was associated with decreased essentialist responding. To the extent that femininity represents culturally accepted stereotypical traits of “femaleness,” based on Eidson and Coley’s (2014) finding that females produce fewer essentialist responses, only the result that femininity predicted decreased essentialist responding – or greater susceptibility to transformation – for political transformations was consistent with our prediction.

**Religiosity.** Religiosity was by far the most significant predictor of participants’ reasoning about social transformations. Further, religiosity was either significantly or marginally significantly associated with increased essentialist responding for male participants’ reasoning about racial transformations in the both condition ($\beta = -.485, p = .075$), participants’ overall reasoning about gender transformations ($\beta = -.28, p = .011$), male participants’ overall reasoning about gender transformations ($\beta = -.338, p = .054$), participants’ reasoning about gender transformations in the both condition ($\beta = -.675, p = .001$), female participants’ reasoning about gender transformations in the physical condition ($\beta = -.422, p = .075$), female participants’ reasoning about gender transformations in the both condition ($\beta = -.654, p = .073$), male participants’ reasoning about gender transformations in the both condition ($\beta = -.730, p = .008$), and male participants’ reasoning about religious transformations in the both condition, $\beta = -.603, p = .055$. In only one instance, male participants’ reasoning about racial transformations in the
characteristic behavioral condition ($\beta = .860$, $p = .017$), religiosity was associated with decreased essentialist responding. In general, these findings are consistent with the notion that religiosity is associated with increased essentialist reasoning.

**Discussion**

The main purpose of Experiment 4 was to examine whether social categories were susceptible to transformations in a manner similar to natural kind categories (e.g., Keil, 1989; Rips, 1989). Further, we wanted to explore whether different types of social categories were more vulnerable to transformations of particular properties which would indicate that such properties are more central to category membership – consistent with the notion of essentialist reasoning.

Our results showed that while participants believe that some degree of transformation is possible for all of the social categories examined, they appear to hold greater essentialist beliefs about racial categories compared to other social groups. In all conditions, participants’ responses indicated that membership in racial categories was less affected by transformations. Just as the results of experiment 3 suggested that the category of gender may be privileged with respect to its inherent potential, it may be that racial categories are privileged with respect to their resistance to transformations.

Although race exhibited less susceptibility to transformation overall, different types of transformations differentially affected membership in social categories. Further, these differential effects seem to mirror the naturalness and cohesiveness dimensions previously reported in social cognition (Haslam, Rothschild, & Ernst, 2000). Race and gender, which have previously been shown to be considered as more natural kind-like categories, are viewed as more susceptible to change when the transformation is to physical properties than changes to
stereotypical behaviors. However, religion and political affiliation – categories which have been shown to be viewed as more cohesive in nature – are rated as more likely to change when the transformation is to characteristic behaviors than to nominal properties. Further, for both sets of categories, the both transformation condition showed no additive effects above and beyond the physical condition for race and gender or the characteristic behavioral condition for religion and political affiliation. This further emphasizes the interpretation that specific transformations differentially affect specific categories by ruling out the notion that simply more property changes result in more category change. Finally, given that transformations to the nominal properties of political groups actually resulted in less change in category membership than changes to less stereotypical characteristic behaviors, there is little evidence that results for gender and political affiliation can be attributed to simply changing more stereotypical properties. While this cannot specifically be ruled out in the case of gender – as greater changes in category membership resulted from changing the more stereotypical physical properties – it seems unlikely that stereotypicality is the sole driving force behind this change given the opposite effect observed for political groups.

While differences in category membership ratings were observed at the group level, we also showed that individual response patterns tend to mirror the group averages. Analyses showed that for religion and political affiliation, individuals were more likely to give essentialist responses in the nominal transformation condition than other conditions and non-essentialist responses in the characteristic behavioral and both transformation conditions compared to the nominal condition. Although not significant, patterns of responses for race and gender trended in the opposite direction with more essentialist responses given in the stereotypical behavioral condition compared to physical or both and more non-essentialist responses in the physical and
both conditions compared to stereotypical behaviors. It should be noted that unlike previous individual response pattern analyses, response patterns in this case were characterized by the magnitude of a single difference score as opposed to a qualitative description of multiple responses. A more sensitive transformation measure may result in significant differences in patterns of responses for race and gender, though these data still trend in the predicted direction. These response patterns further emphasize that categories which are more natural and less cohesive like race and gender are not as susceptible to behavioral transformations as physical ones while categories which are more cohesive and less natural like religion and political affiliation are not as susceptible to nominal transformations as characteristic behavioral ones.

In their previous work, Gelman and Wellman (1991) showed that changing more causally central properties of animals – blood and guts, for instance – led children to believe that the animal had actually shifted category. We may similarly infer that if a change in category membership is believed to have occurred following a transformation, then the properties which were transformed may be more closely causally linked to the essence of the category. If this is the case, then these data may suggest that social categories do not have a general set of essential properties. Rather, social categories which are perceived to be more natural but less cohesive, like race and gender, may be believed to have essences which are more tightly linked to physical, biological, or internal properties. Conversely, social categories which are perceived to be more cohesive and less natural, like religion and political affiliation, may be believed to have essences which are more tightly linked to more behavioral properties. This may predict that social categories which rate highly on both dimensions of essentialism have comparatively more complex essences with links to both physical and behavioral properties. Such categories may presumably be even more resistant to a single transformation than those examined in the current
study as transforming a subset of properties would still leave many essential properties in tact. However, such categories may show greater susceptibility to the both transformation condition in a manner not observed in these data. Overall, this suggests that there may be no single form of social essentialism. Instead, different social categories may be relatively equally essentialized in the average sense, but exhibit differential essentialist effects depending on the perceived underlying nature of that category.

The results of the individual difference measures analyses are much less clear. Measures of motivation to avoid prejudice showed opposing associations with participants ratings about racial transformations. Internal motivation to avoid prejudice was associated with more essentialist responding and preservation of category membership while external motivation to avoid prejudice was associated with less essentialist responding and greater susceptibility to change for male participants’ reasoning about racial transformations in the both condition. While the latter is consistent with Plant and Devine’s (1998) past findings that increased motivation to avoid prejudice should be associated with decreased stereotyping – a proxy for essentialist reasoning – by showing that external motivation to avoid prejudice was associated with a greater willingness to accept transformation, the former is the exact opposite association we would expect. One possible explanation is that externalized pressure to avoid stereotyping leads to greater acceptance of an individual’s ability to change categories while internalized desires to avoid stereotyping may actually reinforce the notion that they cannot. However, this is the first instance in this work in which measures of motivation to avoid prejudice make opposing predictions. Thus, it is unclear if this may simply be an artifact for a particular subset of our participants in a particular condition.
Similar to the results for motivation to avoid prejudice, political affiliation and RWA made opposing predictions about participants’ transformation ratings. For participants reasoning about gender transformation in the physical condition, conservatism was associated with decreased essentialist responding or a greater willingness to accept transformation while RWA was associated with increased essentialist responding or a preservation of category membership for male participants reasoning about racial transformations in the behavioral condition. The latter is consistent with Altemeyer’s (1981; 1996, 2006) past findings that RWA is associated with increased stereotyping and by proxy essentialist reasoning while the former is inconsistent with past work by Napier and Jost (2008) and others showing that conservative ideology is associated with increased stereotyping. This may indicate that these measures are not simply predictive of increased essentialism overall but that various aspects of these constructs are differentially predictive of various aspects of essentialism.

For gender role measures, only femininity was predictive of essentialist responding but, again, in different directions in different conditions. To the extent that femininity represents culturally accepted aspects of “femaleness,” we would have expected based on Eidson and Coley’s (2014) finding that females exhibit less essentialist responding overall that femininity would be associated with decreased essentialist responding. This was only true for females’ responding about political transformations in the nominal transformation condition. In the behavioral racial transformation condition for all participants and males alone, femininity actually predicted greater essentialist responding which was inconsistent with previous findings. This may suggest that femininity specifically or gender roles more generally may be predictive of essentialist responding but in different ways for different categories.
Across categories and conditions, religiosity was by far the most consistent predictor of essentialist responding. Participants who self-reported that religion gave more meaning to their everyday lives were generally less willing to accept changes from one social category to another (i.e., were more essentialist). This is consistent with past findings by Kelemen (2003) which suggested that children in the US – a comparatively more religious cultural environment – were more promiscuous in their teleological bias than children in the UK – a comparatively less religious cultural environment. The teleological bias is similar to the essentialist bias in many ways: it emerges early in development, is promiscuous in children’s reasoning, becomes more selectively applied into adulthood, and re-emerges under time pressure. Thus, these results suggesting that participants who exhibited greater religiosity were more essentialist and less willing to accept social transformations from one category to another closely mirrors the finding that religiosity is associated with increased teleology.

Lastly, although not explicitly designed to test the dual processing account of essentialism, reactions times were collected for each categorization rating to determine if response times might differ depending on transformation condition and category. However, no effects of category, transformation condition, or an interaction between the two were observed. While no relationship was found between categories, conditions, and speed of responding, this does not necessarily discount the dual processing account of essentialism. First, time pressure was not manipulated at all. If responses were solicited under time pressure in future iterations of this experiment, participants may be less willing to accept social transformations than under time delays. Second, it may be that, in general, the wording of this task does not adequately activate Type 2 editing of essentialist responses in such a way as to extract a relationship between essentialist responding and reaction time. A modified version of the task using time pressure or
asking about particular properties of individuals similar to the SAB may be able to establish
relationship between essentialist responding and reaction time.

Experiment 5 will seek to further examine these same questions in a third aspect of
adults’ essentialist reasoning – inductive potential. When added to the results of all previous
experiments, a more complete picture of the extent and disambiguation of adults’ social
essentialism will be determined.

**Experiment 5**

While Experiment 4 focused on the naturalness dimension of essentialism through
investigations of resistance to change, the aim of Experiment 5 was to examine the cohesiveness
dimension by determining the inductive potential of the four social categories used in the
previously outlined experiments. As noted, inductive potential refers to the extent to category
membership is considered informative for making inferences and predictions about category
members and their shared properties. This is a consequence of a belief that category members
each have an essence which gives rise to their category-based properties. Thus, if two individuals
belong to a particular category and are both assumed to have the same underlying category
essence, it is reasonable to assume that they will share similar properties (Gelman, 2003).
Participants’ willingness to make such inferences would suggest essentialist beliefs such that the
members of a category are perceived to be uniform and thus share properties as predicted by the
naturalness dimension of essentialism.

Inductive potential is usually evaluated through a triad task in which participants may see
two individuals each belonging to two different categories and having two different types of
properties (e.g., Birnbaum et al, 2010; Deeb et al, 2011; Gelman, Collman, & Maccoby, 1986;
Hirschfeld, 1996). Then, a target individual is presented sharing one category with each of the
premise individuals, and participants must determine the type of property that the target will have. For instance, Birnbaum et al presented participants with an individual belonging to the category “boy” and “Arabic” and were told that he wanted to be a “Mashitz” when he grew up. Participants also saw a “girl” who was “Jewish” and told she wanted to be a “Nagim” when she grew up. The target individual which participants had to make an inference about was “Arabic” like the first individual but a “girl” like the second individual. Participants then decided which profession she would engage in when she grew up. If participants chose “Mashitz” it was assumed that the a religious category “Arabic” contained more inductive potential and was more informative for making inferences while predicting that she would be a “Nagim” would signal that a gender category “girl” was more informative.

The triad task has been used to examine inductive potential for racial and ethnic categories (e.g. Birnbaum et al, 2010; Hirschfeld, 1996), gender categories (e.g. Gelman, Collman, & Maccoby, 1986), and religious categories (e.g. Diesendruck & haLevi, 2006). In all cases, children ascribe inductive potential to these social groups even in the absence of identifiable physical cues to group membership, with only category labels, and in the presence of perceptual dissimilarity (e.g., an individual labeled as a boy but having the appearance of a girl). However, as previously outlined, adults when they are included in these studies at all often fail to ascribe inductive potential to social categories at all. Instead, they may rely on psychological traits rather than social category membership to make inferences, for example (e.g., Diesendruck & haLevi, 2006).

Still, the triad task has an inherent limitation. Because the task pits the presence of two categories against one another in a forced choice, the results of a triad task can only reveal the relative inductive potential of one category over another; it cannot provide evidence for absolute
inductive potential. Thus, in the example above, if participants ascribe inductive potential to a religious category, it is not clear if this occurs because it is vastly more informative than a gender category or only slightly more informative than a gender category for making inferences.

Further, it may be that participants believe that neither category is particularly useful for making inferences but due to the nature of forced choice responding, the slightly more informative category appears to have significant inductive potential. Similarly, it may be that both categories are believed to have very high inductive potential, but one category is slightly more informative and consistently outweighs the other. The triad task unfortunately cannot distinguish between these scenarios or independently determine each category’s level of inductive potential.

To skirt this problem, Smyth and colleagues (2012; submitted) used a dyadic task in which participants were able to make “yes”/”no” inferences about whether two individuals would share particular properties. Each individual still belonged to multiple categories, but all combinations of those categories were presented such that participants could make an independent judgment as to whether various individuals would share novel properties. By not placing categories in direct competition, this allowed for an absolute measurement of each category’s inductive potential. For instance, participants may see a base picture in which they are told that a child goes to a Catholic church and owns a pet hamster. Participants would also be told that the child had a particular novel property (e.g. “Look at this child. This child is ‘wellant.’”). Participants would then view four possible matches to the original target and decide if those children were also “wellant.” For instance, participants may see another child who goes to a Catholic church and owns a pet hamster and be asked if they thought that this particular child was also “wellant.” Participants might then see a child who goes to a Protestant church and owns a pet hamster and determine if that child was “wellant.” The remaining combinations (a
Catholic, goldfish owner and a Protestant goldfish owner) would also be presented. This allows each category’s inductive potential to be observed in an absolute manner by allowing participants the opportunity to make a yes or no inference about all combinations of category matches. Thus unlike triad tasks in which it is only known whether a category has more inductive potential than another, the dyad task ultimately determines an independent inductive potential score based on the number of inferences made in the presence of each category.

While adults have shown minimal beliefs in the inductive potential of social categories in the past, it is not clear if this is simply due to the limitation of relative measurement in the triad task or if this is actually the result of a developmental shift in essentialist reasoning. The use of the dyadic inference task will be able to reveal the absolute levels of inductive potential for these categories. Thus, the goal of this experiment is to assess the inductive potential for four social categories (gender, race, religion, political affiliation) using a method similar to the dyadic inference task developed by Smyth and colleagues. A base figure belonging to one of each of the critical four critical social categories will be identified as having a novel behavior or physical property similar to the SAB experiments. Then, as in Smyth et al’s task, participants will view all possible combinations of the four social categories, allowing them to make independent inferences about whether individuals share novel properties in the presence of all categories allowing for an absolute measurement of inductive potential for each category.

Based on past induction evidence (e.g., Diesendruck &haLevi, 2006), we might expect that social categories will have relatively low inductive potential for our adult participants based on the developmental shift toward using other types of categories as sources of inference (i.e., psychological characteristics). However, analogous to Experiment 4, Haslam and colleagues’ (2000) dimensions of social essentialism make particular predictions about the informativeness
and inductive potential of each category. For instance, categories like religion and political affiliation were rated highly on the “cohesiveness” factor which refers generally to perceptions of shared group goals and behaviors than race and gender. This may result in religion and political affiliation exhibiting greater inductive potential overall. However, cohesiveness refers to shared behaviors while the naturalness factor generally refers to perceived innateness of traits with the interaction of these two dimensions possibly predicting differential inductive potential based on property. This might predict that religion and political affiliation would exhibit more inductive potential when reasoning about behaviors while race and gender would exhibit more inductive potential when reasoning about physical properties.

Although no time pressure manipulation was employed, the dual processing account of essentialism was examined, as in experiments 1 and 4, by recording the response times for each response. For the purposes of this experiment, we will weight measurements of inductive potential based on the number of categories shared between the target and base individuals and correlate these weighted ratings with reaction times. As in experiment 1 and 4, if essentialism is a fast, automatic T1 bias, we should expect that more essentialist responses – or greater inductive potential – should be given faster than the slower, deliberate less essentialist responses produced by T2 processes.

Finally, as in experiments 2 through 4, we assessed the extent to which individual difference measures are predictive of inductive potential for these social categories.

**Method**

**Participants**
128 participants (64 females) were recruited from Northeastern University’s undergraduate psychology pool. Each participant completed two version of the inference task - one with a novel behavioral one with a novel physical property.

**Materials**

**Inference Task.** Adapted from Smyth and colleagues (2012), participants completed the inference task by viewing an initial base picture which depicted an individual belonging to each of four social categories (i.e., gender: male/female, race: black/white, religion: Christian/Muslim, and political affiliation: Republican/Democrat). For instance, the base individual may have been a black female Christian Democrat. Race was defined by dark or light skin color (black / white) while gender was defined by short or long hair (male / female). Religion was made explicit by a cross or crescent pendant (Christian / Muslim), and political affiliation was made explicit by a red elephant or blue donkey on the individual’s clothes (Democrat / Republican). Either a behavioral (i.e., this individual likes to blicket) or physical (i.e., this individual has wellant cells) novel property was applied to this individual. After viewing the base figure, participants viewed 16 different target figures in random order (see Appendix C for all figures). Each target figure was a possible permutation of the four social categories (e.g., white male Christian Democrat). Participants indicated how likely it was that each target shared the same novel property as the base figure (see figure 13).

Thus, participants reasoned about both novel properties (behavioral, physical) and made inferences about all possible category combinations (gender, race, religion, political affiliation) within subjects. Counterbalancing ensured that all possible combinations of social categories appeared equally as the base across all participants.
Individual Differences. Participants also completed the same battery of individual difference measures used in experiments 2 through 4: motivation to avoid prejudice (internal and external), political affiliation, RWA, gender role (masculinity and femininity as measured by the BSRI), and religiosity.

Procedure

Participants completed both the inference task and individual difference measures on a computer using the Qualtrics survey software (Qualtrics, 2013). Participants were initially given instructions on the structure of the inference task and how to identify the various categories to which a target belonged. Participants began the task by viewing a base individual belong to 4 social categories (race, gender, religion, and political affiliation), and were told that the individual had a novel behavioral or physical property depending on the condition (e.g., “Look at this individual. This individual likes to bicket.”). Participants then received each of the 16 possible targets, one at a time, in random order. For each subsequently presented target figure, participants responded how likely it was that the target (e.g., a black male Christian Republican) would share the novel property with the base figure from 1 (very unlikely) to 9 (very likely). Upon completion of the task, participants received each of the individual difference measures in random order and within each measure received each of the items in random order. Participants then completed a second version of the inference task in which they reasoned about the remaining novel property – either physical or behavioral.

Results

Scoring

In order to calculate the inductive potential for each social category, likelihood ratings for sharing each property were averaged across each of the eight category matches and mismatches
for each social category of interest. For instance, for race, eight targets belonged to the same racial category as the base regardless of other category matches or mismatches. Participants’ ratings that each target would share either a novel behavior or novel physical property were averaged together. Ratings were also averaged for the eight racial mismatches for each property. Match and mismatch scores were calculated separately for each social category (see Table 9).

It should be noted that because each target belonged to four separate social categories, ratings for each target were necessarily used multiple times in separate match or mismatch averages. However, each rating was used an equal four times across the appropriate conditions. Similarly, although match averages for a given social category also include targets which matched on other social categories, the corresponding mismatch ratings for other social categories were also included for the particular social category of interest, thus inherently controlling for the presence or absence of other social categories.

Analyses

**Inductive Potential.** This main aim of this experiment was to determine the extent to which adults ascribe inductive potential to different social categories and whether this differs based on reasoning about different properties or categories. For both behavioral and physical properties, initial paired-samples t-tests showed that match ratings were significantly higher than mismatch ratings for all social categories (all \( ps < .001 \)). This suggests that each social category does exhibit its own inductive potential beyond the influence of other categories. To better observe the interactions of property and category, a difference score was calculated for each participant with mismatch ratings subtracted from match ratings for each category for both properties. The differences scores were subsequently averaged with scores closer to zero representing smaller differences in likelihood of sharing a novel property between targets.
matching in category and those that mismatch – a measure of less inductive potential – and higher scores representing a greater difference in the likelihood of sharing a novel property between targets that match in category and those that mismatch – a measure of higher inductive potential. These difference scores were then entered into a 4 (category: race, gender, religion, political affiliation) by 2 (property: behavioral, physical) repeated measures ANOVA.

**Overall Inductive Potential.** The ANOVA revealed a main effect of property \((F(1,127) = 6.976, p = .009)\) such that inductive potential was higher when reasoning about behaviors \((M = 1.284, SD = .597)\) than physical properties \((M = 1.105, SD = .816)\) and a main effect of category \((F(3,381) = 12.184, p < .001)\) with Tukey pairwise comparisons showing that overall gender \((M = 0.784, SD = 1.161)\) exhibited significantly less inductive potential than race \((M = 1.380, SD = 1.542)\), religion \((M = 1.381, SD = 1.284)\), and political affiliation \((M = 1.232, SD = 1.263)\) (all \(ps < .001\)), and political affiliation exhibited significantly less inductive potential than religion \((p = .050)\), though race did not significantly differ from either religion or political affiliation in inductive potential. This partially supported our prediction that a less cohesive category like gender would exhibit less inductive potential than more cohesive categories like religion or political affiliation, though race – also a less cohesive category – did not differ in inductive potential from more cohesive categories.

**Inductive Potential By Property.** Critically, a category by property interaction was observed \((F(3, 381) = 58.156, p < .001)\). Separate repeated measures ANOVAs were conducted on difference scores for behavioral and physical properties using category as the independent variable (see Figure 14). When reasoning about a novel behavior, a significant effect of category was found \((F(3,381) = 41.716, p < .001)\) with Tukey pairwise comparisons showing that religion \((M = 1.871, SD = 1.130)\) exhibited greater inductive potential than race \((M = 0.857, SD = 1.055)\)
and gender ($M = 0.670, SD = 1.020, ps < .001$), as did political affiliation ($M = 1.736, SD = 1.182, ps < .001$). However, religion and political affiliation did not significantly differ from each other ($p = .239$), nor did race and gender ($p = .118$).

When reasoning about a novel physical property, a significant effect of category was found ($F(3,381) = 22.547, p < .001$) with Tukey pairwise comparisons showing that race ($M = 1.903, SD = 1.763$) exhibited significantly higher inductive potential than gender ($M = 0.897, SD = 1.281$), religion ($M = 0.892, SD = 1.245$), and political affiliation ($M = 0.728, SD = 1.137$) (all $ps < .001$). While gender did not significantly differ in inductive potential from religion ($p = .971$) or political affiliation ($p = .254$), political affiliation exhibited significantly less inductive potential than religion ($p = .039$).

Paired-samples t-tests were also conducted to determine whether inductive potential was greater for a particular property for each category. Significantly for race ($t(127) = 6.88, p < .001$) and marginally for gender ($t(127) = 1.736, p = .085$), greater inductive potential was ascribed when reasoning about physical properties. For religion ($t(127) = 8.221, p < .001$) and political affiliation ($t(127) = 8.147, p < .001$), significantly greater inductive potential was ascribed when reasoning about behavioral properties. These analyses supported our prediction that inductive potential would be differentially applied to categories based on their levels of naturalness and cohesiveness.

**Timing effects.** Although this experiment was designed to primarily investigate other aspects of adults’ social essentialism rather than directly testing the dual-processing account outlined in experiments 1 through 3, reaction times were collected for responses to each target in order to determine whether there was any observable relationship between more essentialist responses and how quickly participants responded. To analyze this, we first weighted
participants’ likelihood rating responses for each target by the number of categories that matched to the base by multiplying the rating by the following factors: 5 for four category matches, 4 for three category matches, 3 for two category matches, 2 for one category match, and 1 for no category matches. Higher weighted ratings represent greater likelihood of sharing a novel property taking into account the number of category matches – more essentialist responses – while lower weighted ratings represent less likelihood of sharing a novel property taking into account the number of category matches – less essentialist responses. Each rating was then correlated with corresponding reaction times for each participant, resulting in 128 individual correlations representing the association between essentialist responding and response time.

The mean correlation ($M_r = -.035, SD = .195$) was then submitted to a one-sample t-test against zero ($t(128) = 2.041, p = .043$) suggesting that these correlations show that essentialist responses are associated with faster response times. This is consistent with the prediction that more essentialist responses – a result of fast, automatic T1 processes – should be produced faster than less essentialist responses – a result of inhibition by slower, more deliberative T2 processes.

**Individual differences.** As in previous experiments, multiple regression analyses were conducted to determine if any relationship could be observed between various individual difference measures and essentialist responding – in this case, differences in match and mismatch likelihood ratings. Analyses were conducted separately for each social category and property but were then repeated separately for male and female participants. The results of all regression analyses are summarized in Table 10 with individual predictors or models having $p < .09$ marked as at least marginally significant.

**Models.** As in previous experiments, models were generally poor at predicting essentialist responding overall. However, taken together, the individual difference measures were
able to predict differences between match and mismatch ratings when all participants were reasoning about a novel behavior for race \( (F(7,120) = 2.417, p = .024, R^2 = .124) \), and female participants’ reasoning about a novel behavior race \( (F(7,56) = 2.622, p = .02, R^2 = .247) \) and male participants’ reasoning about a novel behavior for race separately \( (F(7,56) = 2.388, p = .033, R^2 = .23) \).

**Motivation to Avoid Prejudice (Internal / External).** Bother internal and external motivation to avoid prejudice were associated with smaller differences in likelihood ratings between category matches and mismatches – or less inductive potential.

Internal motivation to avoid prejudice was significantly associated with less inductive potential for both all participants’ reasoning \( (\beta = -.207, p = .037) \) and female participants \( (\beta = -.316, p = .018) \) about a novel behavioral property for race. It was also significantly associated with less inductive potential for female participants’ reasoning about a novel behavior for gender \( (\beta = -.301, p = .031) \) and marginally associated with less inductive potential for female participants’ reasoning about a novel physical property for gender \( (\beta = -.258, p = .079) \). Lastly, internal motivation to avoid prejudice was marginally associated with less inductive potential for female participants’ reasoning about a novel behavior for religion \( (\beta = -.260, p = .064) \).

External motivation to avoid prejudice was a significant predictor in one instance and was associated with less inductive potential for male participants’ reasoning about a novel behavior for political affiliation \( (\beta = -.289, p = .031) \).

**Political Affiliation.** Political affiliation was significantly associated with less inductive potential for male participants’ reasoning about a novel behavior \( (\beta = -.344, p = .046) \) and physical property of gender \( (\beta = -.344, p = .049) \). Political affiliation was also marginally associated with female participants’ reasoning about a novel behavior for race \( (\beta = -.299, p = .064) \).
.080) and participants’ reasoning about a novel physical property of gender ($\beta = -0.219, p = 0.067$).

In all cases, conservatism was associated with less inductive potential counter to previous predictions that conservative political ideology would be associated with greater essentialist responding.

**RWA.** In contrast to political affiliation, RWA was associated with greater essentialist responding in many cases. RWA was significantly associated with participants’ reasoning about and novel physical property for race ($\beta = 0.311, p = 0.018$) and male participants’ reasoning about a novel behavior for race ($\beta = 0.546, p = 0.006$) and gender ($\beta = 0.445, p = 0.034$). It was also marginally associated with participants’ reasoning about a novel behavior for race ($\beta = 0.240, p = 0.060$) and gender ($\beta = 0.249, p = 0.060$) and male participants’ reasoning about a novel physical property of religion ($\beta = 0.414, p = 0.052$).

**Gender Roles (Masculinity / Femininity).** Both masculinity and femininity showed mixed patterns of association. Masculinity was significantly associate with less inductive potential for female participants’ reasoning about a novel behavior of race ($\beta = -0.282, p = 0.030$) and greater inductive potential for their reasoning about a novel behavior of religion ($\beta = 0.301, p = 0.029$). Masculinity was also marginally associated with greater inductive potential for female participants’ reasoning about a novel physical property of race ($\beta = 0.264, p = 0.056$).

Femininity only showed associations for reasoning about novel behaviors of political groups. It was significantly associated with greater inductive potential for male participants ($\beta = 0.318, p = 0.014$) but marginally associated with less inductive potential for female participants ($\beta = -0.258, p = 0.064$).

In general, we predicted that masculinity should be associated with greater essentialist responding – or greater inductive potential in this case – and that femininity should be associated
with less essentialist responding. In both cases, results were mixed with half or more associations in the expected directions, with the remaining associations exhibiting unexpected patterns.

Religiosity. Religiosity was associated with significantly less inductive potential for both all participants’ (β = -.231, p = .034), generally, and male participants’ (β = -.597, p = .001), specifically, reasoning about a novel behavior of race. It was also marginally associated with less inductive potential for all participants’ (β = -.205, p = .069), generally, and male participants’ (β = -.320, p = .074), specifically, reasoning about a novel behavior of gender and male participants’ reasoning about a novel physical property of political groups (β = -.327, p = .069). However, for participants’ reasoning about a novel behavior of religion, religiosity was marginally associated with greater inductive potential (β = .192, p = .089). Thus, our prediction that religiosity would be associated with greater essentialism was supported only for reasoning about religious categories, but opposing associations were found for other categories.

Conclusion. As in previous experiments, there was no single individual difference measure which was able to predict participants’ essentialist responding across categories or conditions. Most of the individual difference measures show expected patterns of association with essentialist responding with a few exceptions. However, these results lend further evidence to the possibility that multiple factors may be involved in regulating essentialist responding under various conditions.

Discussion

The main goal of this experiment was to determine whether adults ascribed inductive potential to social categories and whether this might differ based on a novel behavior or physical property. Specifically, we predicted that more cohesive categories like religion and political
affiliation might show greater inductive potential overall than less cohesive categories like race or gender. However, we predicted that categories’ cohesiveness and naturalness would interact with the property being reasoned about with more cohesive categories exhibiting more inductive potential when reasoning about behavioral properties but more natural categories exhibiting more inductive potential when reasoning about physical properties.

Initial results showed that regardless of category or property type, participants made higher ratings for category matches than mismatches suggesting that overall participants do ascribe inductive potential to the social categories that were investigated. That is, for each category of interest, even though mismatches trials may actually match on other categories, participants rated that individuals in each category’s match trials would be more likely to share novel properties. Further, more cohesive categories like religion and political affiliation exhibited greater inductive potential than the less cohesive category of gender – but not race. This generally aligned with our prediction that cohesiveness may predict overall inductive potential, though the category of race – though not perceived as particularly cohesive – was still as informative for making predictions as more cohesive categories.

Critically, however, participants rated individuals as more likely to share a novel behavior when they shared the same religion or political affiliation than the same race or gender. Conversely, participants rated individuals as more likely to share a novel physical property when they shared the same race but not other categories. Further, inductive potential as measured by the differences in likelihood ratings between match and mismatch trials was greater for religious and political affiliation when reasoning about behaviors than when reasoning about physical properties while the reverse was true for race and gender - greater inductive potential was afforded when reasoning about physical properties than behaviors.
Taken together, these results suggest that overall each of the social categories investigated exhibited some amount of inductive potential. However, the amount of inductive potential ascribed to each category was contingent on the type of novel property being reasoned about. Specifically, religion and political affiliation - categories which are perceived to be more cohesive – showed the greatest inductive potential when reasoning about behaviors and greater inductive potential about behaviors than physical properties. Race – a category which is perceived to be more natural – exhibited the greatest inductive potential when reasoning about physical properties. Further, race and gender both exhibited greater inductive potential when reasoning about physical properties than behaviors – though the difference in inductive potential between reasoning about physical properties and behaviors was much greater for race than for gender. This confirms our original hypothesis that the nature of the social category and property being reasoned about may interact to determine inductive potential.

Although inductive potential as measured by the difference in match and mismatch ratings was significantly greater than zero, it was still relatively low overall with mean differences just over one. Still, the likelihood of shared property for matches was greater than the scale’s midpoint for all categories for both properties. Ratings for mismatches were also relatively high, though all averages were less than the scale’s midpoint for all categories for both properties. This may inherently be due to the fact that among mismatches, categories other than the one of interest can match to the base. This raises mismatch ratings overall, though these are still controlled by the various non-matches included in the match ratings and through the difference scores. Given this, we can conclude that overall adults do display inductive potential for the social categories investigated, though this potential differs based on category and property type.
Although not explicitly designed to test the essentialist dual processing claim with a speeded paradigm as in previous experiments, reaction times gathered in this experiment allowed us to test whether more essentialist responses – in this case a greater likelihood that two individuals would share a novel property taking into account the number of category matches between the target and base – would be associated with faster responding. Indeed, the average correlation between weighted likelihood ratings and response times was significantly different from zero and showed that essentialist responding was associated with faster response times. Though the strength of the average correlation was not particularly strong, these effects are consistent with the dual processing essentialist claim. If essentialism is a fast, automatic bias which can quickly produce essentialist responses – which then may or may not be edited by slower T2 processes – then more essentialist responses or ratings should be made faster than less essentialist ones. If a more forced choice version of the inference task was designed similar to the forced choices required in the SAB task in Experiment 1, stronger timing effects may be observed.

Lastly, as in previous experiments, we attempted to determine whether essentialist responding could be predicted by individual difference measures to determine what processes may be involved in suppressing essentialist responding. While the regression models were again relatively poor at predicating essentialist responding, they were able to predict participants’ inductive potential for race as evidenced by the significant models for all participants and males and females separately when reasoning about behaviors. More interestingly, most of the individual difference measures were predictive across various category, property, and participant gender conditions.
For instance, both internal and external motivation to avoid prejudice were predictive of inductive potential in various conditions for each of the four social categories with increased motivation associated with less inductive potential. These relationships are consistent with past findings like those by Plant and Devine (1998) which show that increased motivation to avoid prejudice is associated with decreased stereotyping. If participants either hold personal beliefs that stereotyping is wrong or feel external pressure to avoid stereotyping, then we would indeed expect that they would hesitate to endorse the notion that individuals share a property based on their social category membership alone.

Similarly, increased RWA was shown to be predictive of increased inductive potential for race, gender, and religion for both physical and behavioral properties in certain conditions. These results are consistent with past findings showing that RWA is associated with increased stereotyping and prejudicial attitudes (Altemeyer, 1981; 1996, 2006). This is expected given that inductive potential assumes that shared essences point to shared properties similar to how stereotypes assume that category membership points to a culturally specified set of properties.

Interestingly, political affiliation and religiosity are both predictive of inductive potential in the opposite expected direction. Political affiliation – specifically conservatism – has been associated with increased stereotyping similar to RWA (e.g., Napier & Jost, 2008; Sidanius, Pratto, & Bobo, 1996) predicting that conservatives should also exhibit greater inductive potential. However, regression analyses showed that the opposite was true; participants who were more liberal exhibited greater inductive potential in most cases.

Similarly, while we might expect from the results of experiment 4 and past findings from Kelemen (2003) that religiosity should be associated with increased essentialist responding, that is only true of participants reasoning about religious behaviors. In all other cases, increased
religiosity was associated with decreased inductive potential. This could signal that when individual difference measures are specifically relevant or not to a particular social category there are differentially predictive relationships.

While many of the individual difference measures are predictive of inductive potential in the expected directions based on past literature, political affiliation and religiosity notably are not. Further, there is not a single measure that is ubiquitously predictive of essentialist responding in this or other experiments. This provides further evidence that effortful suppression of the essentialist bias may involve many complex factors that depend on the demands of a task or the category being reasoned about.

Overall, this experiment shows that like other aspects of essentialist reasoning, adult participants do exhibit inductive potential for a range of social categories, but critically inductive potential is selectively attenuated based on property. Specifically, more cohesive categories are likely to exhibit inductive potential when reasoning about behaviors and more natural categories are more likely to exhibit inductive potential when reasoning about physical properties. However, future investigations into adults’ social induction may show more across-the-board generalizations by category and property if cognitive load or time pressure paradigms such as those used by Eidson and Coley (2014) or Experiments 2 and 3 are employed. The property effects that we see here are reminiscent of selective response patterns seen in past work on inherent potential (e.g., Taylor et al, 2009) where adults gave essentialist responses about physical properties of gender but not about physical properties. However, Eidson and Coley as well as Experiments 2 and 3 showed that adults became more promiscuous in their essentialism under time pressure. Thus, the selective patterns that we see here may similarly be a result of suppressing certain essentialist impulses. However, it is also possible that the dyadic inference
task did not trigger inhibitory mechanisms involved in suppressing essentialist responding such that these patterns represent adults’ true degree of essentialism with regards to inductive potential.

General Discussion

Each of the preceding experiments set out to generally assess the state of adults’ social essentialism. More specifically, we wanted to test and obtain evidence for the claim that essentialism can be viewed in a dual processing framework, determine whether essentialist responding could be predicted by individual differences, and obtain evidence of adults’ social essentialism for multiple aspects of essentialist reasoning and for multiple social categories. The following sections detail the outcome of each of the previously outlined specific aims.

Aim 1: Test the dual processing claim that essentialism can be thought of as a T1 bias.

Eidson and Coley’s (2014) initial work on gender essentialism, found that essentialist responding was greater for participants placed under time pressure compared to those who responded under a time delay. Based on these findings, they posited that essentialism is not just a cognitive bias evident in early childhood but is instead is an intuitive way of reasoning that is active and persists into adulthood. Further, because essentialist reasoning was more prevalent under time pressure, the essentialist bias may be thought of as a fast, automatic Type 1 process which can be overridden by slower and more deliberative cognitive processes. We set out to test this claim in these experiments.

Experiment 1 initially attempted to provide analogous results to Eidson and Coley’s time pressure findings by using a cognitive loading paradigm. While time pressure limits temporal access to various editing mechanisms which may suppress essentialist responding, tapping
cognitive and attentional resources which would otherwise be used for editing essentialist responding for use in a separate task could accomplish the same goal. However, rates of essentialist responding must also be taken together with timing evidence to accurately depict the full range of participants’ responses.

Although we attempted to increase essentialist responding by having participants hold a list of non-words in their memory while completing a gendered SAB task, the cognitive load manipulation was ultimately unsuccessful at replicating the same increase in essentialist responding observed under time pressure. As previously discussed, this may have been due to the choice of a linguistic loading task rather than a non-linguistic one. Overall non-word recall was slightly lower for the loading condition, suggesting that participants may have decreased cognitive load by discarding some non-words from memory altogether.

While the cognitive load manipulation was unable to produce an increase in essentialist responding over baseline, we also examined participants’ self-paced reaction times for essentialist and non-essentialist responses. For reasoning about both behavioral and physical aspects of gender, participants gave their essentialist responses faster than their non-essentialist ones. This is consistent with the notion that essentialism is a fast, automatic Type 1 process. When essentialist responses were not suppressed by Type 2 editing mechanisms, they were ultimately produced faster providing evidence for a dual processing account of essentialist reasoning.

Experiments 2 and 3 returned to the speeded SAB paradigm used by Eidson and Coley (2014) using within-subjects timing manipulations to determine whether individual participants would increase their essentialist responding under time pressure for social categories. Experiment 2 focused on gender essentialism alone while Experiment 3 examined racial, gender,
religious, and political categories. If essentialist responding, especially for behaviors, increased under time pressure, it would suggest that not only do participants who give fast responses exhibit higher rates of essentialism but that an individual can be pushed from one pattern of reasoning to another by virtue of restricting access to slower, more deliberate reasoning processes.

In both experiments, time pressure increased essentialist responding about behaviors. This was true of gender categories alone in Experiment 2, but also of racial, religious, and political categories in Experiment 3. Further, these increases were not merely due to a small overall group increase in essentialist responding but also due to individual participants giving consistently more essentialist sets of responses under time pressure compared to a delay. While it should be noted that in Experiment 3, other than for gender categories, essentialist responding about behaviors was overall low, the more realistic nature of the SAB task (i.e., real-world adoption scenario) compared to those used in past work (i.e., a child growing up on an island) (e.g., Taylor, Rhodes, & Gelman, 2009) and the use of abstract properties instead of category-specific stereotypical properties may have contributed to decreased essentialist responding. Still, increases in essentialist responding about behaviors was reliable at both the group level and the level of individual response patterns. Taken together, this suggests that when participants are forced to quickly make predictions about individuals belonging to the particular social categories examined participants are more likely to resort to an essentialist mode of reasoning. However, when they are given time to think about and edit their responses, participants become selective in their reasoning – still providing essentialist responses about physical properties, but exhibiting minimal essentialist responding about behavioral properties. The former is indicative of
essentialist reasoning as a Type 1 process while the latter is consistent with Type 2 editing of essentialism.

Experiments 4 and 5 were not designed to explicitly test the dual processing framework of essentialism. However, as in Experiment 1, they provided the opportunity to observe participants’ self-paced responding in a social transformation task and a social inference task. If essentialism is a fast and automatic bias, we would have expected that essentialist responding should be associated with faster response times. We examined this in Experiment 4 by comparing conditions in which categories were less susceptible to transformation compared to those that resulted in greater changes in category membership. However, we did not find any differences in reaction times based on transformation condition or category. This may be because the responses given in the transformation task are all indicative of true Type 1 essentialist responding with responses given equally fast. However, it may also simply be that the transformation task is not adequately sensitive to time pressure differences given that each participants only made 8 responses total which were then re-calculated into 4 reaction time differences – one for each category. This may have resulted in insufficient power to detect differences in reaction time among the different conditions and categories.

In Experiment 5, we weighted participants’ ratings for the likelihood that individuals of different categories would share a novel property by the number of categories that the two individuals shared. We then correlated these weighted ratings with reaction times for each person and compared the average correlation across all participants to zero to determine if greater likelihoods of sharing a novel property based on category membership – a more essentialist response – would be associated with faster response times. Indeed, while the average correlation itself was small, greater likelihood ratings were associated with faster response times. This
suggests that when the reasoning system is outputting a more essentialist response – in this case determining that two individuals who share category membership should share a novel property – that response is given faster than less essentialist responses. Again, this is consistent with a dual processing framework of essentialism.

Taken together, these findings have two important implications. First, they provide strong evidence for the dual processing account of essentialism. It is increasingly clear that we must think of essentialism as a fast and automatic bias which, while subject to editing by Type 2 processes, continually influences the reasoning system. Further, it seems likely that this is true of multiple aspects of essentialist reasoning, not just inherent potential as measured by switched-at-birth tasks, though it is not clear quite to what extent this may be true. Future work will have to determine whether time pressure will create different patterns in reasoning for transformation tasks, for instance.

The second important implication is that in light of these findings we must reconsider the ways in which we understand the developmental progression of essentialist reasoning. Work by Taylor, Rhodes, and Gelman (2009), Keil (1989), and Diesendruck and haLevi (2006) among others, have all shown that essentialist reasoning either decreases or becomes more selective for inherent potential, resistance to change, and inductive potential, respectively, as children age into adolescence and young adulthood. However, the findings presented here strongly suggest that this developmental trend is only a progressive masking of essentialist reasoning. We might expect that were these tasks modified with time pressure or cognitive load paradigms children would exhibit similar increases in essentialist responding. It is also likely that essentialist reasoning is never replaced. That is, well into later adult development, it is reasonable to assume
that the essentialist bias continues to influence reasoning. Future work in either developmental direction will help to elucidate these possibilities.

Aim 2: Ascertain what aspects of T2 processing may be involved in limiting social essentialism.

The second goal of the current work was to identify and test relationships between various individual differences and essentialist responding. Starting from the assumption that essentialism can be viewed as a fast and automatic Type 1 process (Eidson & Coley, 2014), we set out to determine what sorts of constructs might be implicated in the Type 2 editing of essentialist reasoning. We first identified three major classes of individual differences which may be associated with essentialist responding: conceptual factors, social factors, and executive function.

Conceptual predictors of social essentialism. For our purposes, we focused on the conceptual factor of willingness to sex-type properties as a form of belief in variability. As previously outlined, to the extent that a category and its members are believed to have an underlying essence which gives rise to category-based properties, then presumably all category members should have a high probability of exhibiting those category-based properties resulting in very little category variability (e.g., Rhodes, Gelman, & Brickman, 2008; Rhodes & Brickman, 2010; Shtulman & Shulz, 2008). Thus, those who instead believe in category variability may exhibit less essentialist beliefs.

However, it appears that beliefs about sex-typing may actually be related to the suppression of essentialist responding. In Experiment 2, we found that willingness to sex-type behaviors – or say that more men or women should exhibit a property than the opposing gender –
were more likely to show increased change under time pressure, though only for male participants. This at least suggests that those who sex-type show increased essentialist responding under time pressure compared to a time delay somewhat comparable to our prediction that they would exhibit increased essentialist responding more generally. However, this also suggests that those who sex-type behaviors are more susceptible to the time pressure manipulation’s ability to increase essentialist responding than those who do not. Those who are unwilling to sex-type – or endorse the notion that men and women should be equally likely to exhibit a particular behavior or physical trait – instead are more likely to exhibit consistent essentialist responding under different timing conditions.

Due to methodological restrictions, we were unable to incorporate this particular individual difference measure into subsequent experiments. As it was imagined, it used participants’ beliefs about the distribution of particular properties used in the switched-at-birth task to predict essentialist responding about gender. However, as Experiments 3 through 5 did not necessarily use or ask participants to make decisions about specific stereotypical properties, this particular paradigm was not appropriate for use in those instances. Future investigations of relations between belief in variability may benefit from generalizing the task for use with other social categories and essentialist tasks. Such a task may be more sensitive to relations between belief in variability and essentialist responding more generally, but may also further elucidate the observed relationship between these beliefs and suppression of essentialist responding at baseline.

**Social predictors of social essentialism.** We also wanted to determine if various social factors might predict essentialist responding in our tasks.
Motivation to avoid prejudice. Based on the previously outlined work by Plant and Devine (1998) and Dunton and Fazio (1997), among others, we expected that motivation to avoid prejudice would be associated with the suppression of essentialist reasoning such that once access to these motivations was restricted via time pressure, essentialist responding would increase. However, results were generally mixed with two relatively distinct patterns of association with internal versus external motivations to avoid prejudice.

For instance, internal motivation to avoid prejudice was associated with decreased changes in essentialist responding about gender in Experiment 2 but increased change in essentialist responding about religion in Experiment 3. The former finding was counter to our prediction that restricting access to Type 2 processes should show positive associations between motivation to avoid prejudice and degree of change in essentialist responding while the latter was the expected pattern. Further, in many cases, internal motivation to avoid prejudice was associated with essentialist responding in general: female participants’ increased essentialist responding about gender in Experiment 3, male participants’ beliefs about racial categories’ resistance to transformation in Experiment 4, and decreased inductive potential for race, gender, and religion for both novel behaviors and physical properties. Since, the direction of these associations often differs by category as well as by task, it may be that there are different abstract aspects of internal motivation to avoid prejudice which affect reasoning differently for different categories or aspects of essentialist reasoning. However, it is also possible that, although we used a single measure of motivation to avoid prejudice with a single prediction for all categories based on past work, that individuals may have different motivations – or lack thereof – to avoid prejudice for different types of categories. Motivations to avoid sexism may not equate to avoiding stereotypical judgments about religious categories or vice versa, for instance.
However, in contrast to these patterns, external motivation was not associated with suppression of essentialism in Experiments 2 or 3 but was instead associated with decreased essentialist responding in a few instances – willingness to accept racial and political transformations in Experiment 4 and less inductive potential for political categories in Experiment 5. As with internal motivation to avoid prejudice, these associations differ by category and task. Further, the associations between essentialist responding are different by category and task for external and internal motivation to avoid prejudice. This suggests that each may have independent sets of pressures which may affect essentialist reasoning about different categories in different ways. It also appears that only personal, internal motivations to avoid stereotyping are related to suppressing baseline essentialism. This suggests that any attempt to use motivation to avoid prejudice should always consider internal and external motivations separately with regard to assessing individual differences in essentialist responding.

**Political Affiliation.** We also investigated two factors associated with political ideology. The first was political affiliation. Conservative political ideology (e.g., Napier & Jost, 2008; Sidanius, Pratto, & Bobo, 1996) has been associated with both increased stereotyping and prejudicial attitudes leading us to predict that political affiliation might be associated with increased essentialist reasoning. However, we found that conservatism was associated with decreased essentialist responding. In both Experiment 2 and 3, conservatism was associated with less degree of change in essentialist responding between the two timing conditions, thus being less likely to show increased essentialism due to the timing condition. In Experiment 4, when reasoning about gender transformations, conservative political ideology was also associated with decreased essentialism with more conservative participants being more willing to accept
transformation; in Experiment 5, conservative ideology was again associated with decreased essentialism showing less inductive potential for racial and gender categories.

We did, however, notice an interesting gender difference with regard to political affiliation’s association to essentialist responding. In Experiment 2 more conservative female participants tended to show increased essentialist responding, but the opposite was true for male participants. The former association is consistent with past work showing that conservative ideology is associated with stereotyping while the latter is not. Further, based on past work, there was no particular reason to expect such opposing associations based on gender. At the least, it is clear that political affiliation may interact with other factors like participant gender. Yet, while political affiliation was certainly predictive of essentialism in many instances, these findings suggest that we must be cautious of drawing parallels between other work on stereotyping and essentialist reasoning. Past work had suggested that conservatism should be associated with increased stereotyping, yet it was often associated with decreased essentialist reasoning across tasks. Thus, while associations with stereotyping may implicate a possible predictor of essentialist reasoning, it may not always be in an expected way.

**RWA.** Similar to political affiliation, RWA has been shown to be associated with both increased stereotyping and prejudicial attitudes (Altemeyer, 1981; 1996; 2006) leading us to originally predict that RWA should be associated with increased essentialist responding. While RWA was not generally predictive of essentialist responding in the switched-at-birth experiments, RWA was predictive of male participants’ essentialist responding in Experiment 4 such that they were less willing to accept racial transformations. In Experiment 5, RWA was associated with greater essentialist responding such that participants with greater self-reported RWA were more likely to attribute greater inductive potential to racial, gender, and religious
categories for both novel behaviors and physical properties. In Experiment 3, RWA was also associated with increased change in essentialist responding when reasoning about race such that those with higher RWA exhibit greater essentialist responding under time pressure compared to under time delay, somewhat consistent with the notion that RWA should be associated with greater essentialist responding more generally. Thus, taken together, RWA was generally associated with increased essentialist responding for all aspects of essentialism that were examined as predicted. While there were many instances in which RWA was not predictive, it may be a candidate for increased scrutiny for predicting essentialist responding more broadly.

**Gender roles.** The third social factor that we considered was gender role as measured by the BSRI (Bem, 1978). Based on Eidson and Coley’s (2014) initial work, we expected that masculinity might be associated with increased essentialist responding, but femininity might be associated with decreased essentialist responding. In general, our results supported these predictions.

For instance, more masculine participants tended to show more essentialist responding about gender in Experiment 2 as well as a greater degree of change in essentialist responding. The former association is certainly in line with the prediction that masculinity should be associated with greater essentialism, and the latter at least suggest that masculine individuals exhibited more essentialist responding under time pressure than they did under time delay. The same was true in Experiment 3 with masculinity associated with greater essentialist responding for both race and gender and a greater degree of change in essentialist responding when reasoning about race. Masculinity was also related to essentialist responding in Experiment 5 with masculine male participants showing greater inductive potential for religious categories, while masculine female participants showed mixed associations with less inductive potential for
racial categories when reasoning about behaviors but increased inductive potential when reasoning about physical properties. In general, these findings are consistent with the notion that to the extent that masculinity taps some culturally accepted aspects of “maleness,” and males may generally provide more essentialist responses, masculinity should be associated with increased essentialism.

On the other hand, femininity was associated with less essentialist responding about religion in Experiment 3 as well as less degree of change when reasoning about race. This generally aligns with the prediction that femininity should be associated with less essentialist responding given that female participants have been observed to show the same pattern. However, results were much more mixed for Experiments 4 and 5. In Experiment 4, feminine male participants were more likely to reject racial transformations as causing changes in category membership – an essentialist response – while feminine female participants were more likely to accept political transformations as changing category membership – a non-essentialist response. Similarly, in Experiment 5, femininity was associated with decreased inductive potential for females when reasoning about a novel behavioral property for political groups – a non-essentialist response – but increased inductive potential for males – an essentialist response.

While femininity was associated with decreased essentialist responding more often than not, as we would expect, such an association was certainly not universal.

Given that gender role provides a separate, orthogonal contrast to gender alone (i.e., people have a particular gender and a gender role which may or may not align with said gender), the two should always be considered in tandem when trying to predict essentialist responding. As we saw in Experiment 4 and 5, feminine males showed opposing associations with essentialist responding to feminine females with the former actually exhibiting greater essentialist
responding, though feminine participants in general showed less essentialist responding in Experiment 3.

Religiosity. The final social factor that we examined was religiosity. As previously discussed, Kelemen (2003) found that children living in less religious environments were more selective in their endorsement of teleological explanations of natural phenomena than children living in more religious environments. Given the similarities between teleology and essentialism as cognitive biases – both emerge early in development, involve a developmental increase in selectivity, and both seem to re-emerge under appropriate conditions in adults – we predicted that religiosity may be associated with increased essentialism. Indeed, of all individual differences examined throughout the presented experiments, religiosity was perhaps the most ubiquitous in predicting essentialist responding across conditions and experiments.

In Experiment 3, religiosity was associated with greater essentialist responding for males reasoning about religion as well as a greater degree of change in essentialist responding for males’ reasoning about race. In Experiment 4, religiosity was associated with greater essentialism such that participants were less willing to accept social transformations for racial, gender, and religious categories. Conversely, for Experiment 5, religiosity was generally associated with decreased essentialism such that participants exhibited less inductive potential for race, gender, and political categories but increased inductive potential for religion.

Thus, religiosity showed expected patterns of association with the inherent potential of categories and their resistance to change but the opposite pattern of association with respect to inductive potential except when participants were reasoning about religion. As with other measures, this suggests that while religiosity may be a relatively good predictor of essentialist responding overall, it may not always predict essentialism in the expected direction. Rather, it
appears that various individual differences or motivations related to essentialism may occasionally make differential predictions based on the aspect of essentialist reasoning involved. Thus, as a field of study we must begin to consider that evidence for one particular aspect of essentialist reasoning should not necessarily be taken as evidence for essentialist reasoning as a whole.

**Executive function as a predictor of social essentialism.** Finally, we also examined executive function as a general inhibitory mechanism which may suppress essentialist responding. If we assume that essentialism is a fast, automatic, Type 1 process, regardless of what factors may motivate suppression of essentialism, our general ability to inhibit our gut responses at all may be indicative of rates of essentialist responding. Based on Kelemen and Rosset (2009)’s finding that performance on the Stroop task predicted the extent of participants’ teleological explanations when they were placed under time pressure, we expected that executive function may be associated with less degree of change in essentialist responding between timing conditions. However, executive function measures were not predictive of essentialist responding or degree of change in responding. This, along with the myriad relationships observed between other individual differences and essentialist responding may suggest that suppression of essentialism in adults is not simply due to a general inhibitory mechanism. Rather suppression may be motivated by certain aspects of explicit knowledge, social pressure, or ideologies.

**Conclusions.** Taking these results together, the preceding experiments have outlined numerous relationships between individual differences and essentialist responding. Many of said relationships follow expected patterns based on past work on essentialism and in other domains. However, in many cases, these relationships were the opposite of what we might have expected – especially in the case of political affiliation. Further, some individual differences predicted
opposing relationships with essentialism depending on the category and particular aspect of essentialist reasoning being examined. This suggests that the interactions between various individual differences and essentialism or its suppression in adults involves an incredibly complex set of interactions.

Depending on the reasoning context – either based on category, property, gender of a participant, or some other factors like age (e.g., Taylor, Rhodes, & Gelman, 2009) or culture (e.g., Birnbaum et al, 2001; Smyth et al, submitted) – one could posit that various triggers from these individual differences could signal Type 2 processes to inhibit essentialist responding. For example, if I feel external pressure to avoid making stereotypical judgments about social groups and then am asked to reason about whether members of a particular category share certain properties or should be grouped together, this may signal a red flag to Type 2 processes to inhibit my intuitive essentialist response. However, if I feel no such external pressure, then Type 2 processes may not be engaged and my intuitive, essentialist response would be allowed as behavioral output.

Further, various personality traits like levels of RWA may help to account for why some individuals exhibit strong levels of essentialist responding and why some, even after being placed under time pressure, simply exhibit very little essentialism. This raises questions as to whether this is a difference which is mostly present in adults or if such differences in general essentialism are prevalent in younger children. Future developmental work on essentialism may benefit from increased scrutiny of individual differences and their changes over time.

**Aim 3:** Determine the extent of adults’ essentialist reasoning.
The third goal of this work was to examine the state and extent of adults’ social essentialist reasoning. We did this by both examining multiple social categories – race, gender, religion, and political affiliation – as well as by using multiple essentialist tasks to examine inherent potential, resistance to transformation, and inductive potential. Haslam et al.’s two orthogonal dimensions of essentialism – naturalness and cohesiveness – made strong predictions for adults’ reasoning in each essentialist task. In general, we expected more natural categories to have more inherent potential as measured by the SAB tasks in Experiments 2 and 3, overall greater resistance to change as measured by the transformation task in Experiment 4, and to hold greater inductive potential for reasoning about physical properties in Experiment 5. Further, we expected that more cohesive categories would hold greater inductive potential for reasoning about behaviors. Overall, we found that our data closely followed these expected patterns.

**Inherent potential.** Experiments 2 and 3 both showed that adult participants’ hold underlying beliefs about the inherent potential of social categories as evidenced by both generally high essentialist responding about physical properties and increases in essentialist responding about behaviors under time pressure in the SAB for all categories. As predicted by its perceived naturalness, participants were generally more essentialist about gender than other categories. Though Haslam et al.’s dimensions of essentialism might have also predicted greater essentialist responding about race, it may be that gender is simply privileged in terms of inherent potential compared to other social categories, but further investigation of other natural categories will be necessary to determine this.

Taken together, these results on the state of inherent potential for adults’ social categories refutes the notion that adults become selective in their application of this particular aspect of essentialist reasoning (e.g., Taylor, 1996; Taylor, Rhodes, & Gelman, 2009). Rather, this
particular aspect of essentialist reasoning is simply suppressed over the course of development. Further, while past studies have only shown inherent potential for the category of gender, Experiment 3 showed that adults exhibit, to a lesser extent, increasing inherent potential for a range of social categories. Future iterations of switched-at-birth tasks should consider examining other age- or culturally-appropriate social categories developmentally to determine the trajectory of essentialist reasoning for such categories.

**Resistance to change.** Experiment 4 was a first investigation into adults’ reasoning about social transformations. As with inherent potential, we expected that more natural categories should generally be more resistant to change than less natural ones. Indeed, race was generally less susceptible to transformation than other categories. Although we also expected gender to be more resilient, it may be that race is also privileged with regard to social transformations. However, as with gender and inherent potential, other natural categories will have to be examined in a social transformation task before this can be definitively determined.

Perhaps more importantly, we found that the types of transformations that most affected social category membership depending on categories’ perceived naturalness and cohesiveness. More natural categories were more affected by physical transformations than behavioral ones, and more cohesive categories were more affected by transformations of characteristic behaviors than nominal properties. Critically, the both condition showed no increase over the transformation – either physical or characteristic behaviors – that affected each category the most suggesting that these patterns of susceptibility were genuine.

Given that Gelman and Wellman (1991) found that transformations targeting the more causally central properties resulted in category change while Keil (1989) and Rips (1989) found that transformations to properties less tied to the category essence did not necessarily result in
change, these results may suggest that the underlying essence of more natural categories like race and gender may be perceived to be fundamentally different from the underlying essence of more cohesive categories like religion and political affiliation. If this is the case, it could possibly account for differences in levels of essentialism across various essentialist tasks. For instance, if religion is believed to have an essence characterized by particular behaviors or religious practices, it would be unsurprising that it might elicit comparatively less essentialist responding in a switched-at-birth task than an inference task. Future work examining these different aspects of essentialist reasoning must consider that essentialism may not be an all-or-none phenomenon. Instead, each social category, perhaps dependent on levels of naturalness and cohesiveness, may exhibit unique levels of each aspect of essentialist reasoning.

**Inductive potential.** Recall that in Experiment 5, we expected similar patterns of results to Experiment 4 with more inductive potential for race and gender than religion or political affiliation when reasoning about a novel physical property, but more inductive potential for religion and political affiliation than race or gender when reasoning about a novel property based on levels of naturalness and cohesiveness, respectively. Indeed, we saw strikingly similar patterns of results for inductive potential as we did for resistance to transformation. Inductive potential followed the pattern predicted by naturalness and cohesiveness precisely. Further, all categories exhibited inductive potential which is counter to previous findings on adults’ social induction which suggested that adults may have been unwilling to make inferences based on social categories at all (e.g., Diesendruck & haLevi, 2006). This may also be partially attributable to our use of a dyadic induction task which allowed us to assess absolute levels of inductive potential instead of the relative levels usually assessed by triad tasks.
As with social transformations, adults’ do exhibit essentialist patterns of reasoning when it comes to social induction. Further, these patterns are predicted by a category’s perceived naturalness and cohesiveness. Again, this may explain inconsistencies for various categories’ levels of essentialist responding across multiple tasks depending on the types of properties being reasoned about. It also makes it clear that while adults’ social induction may be selective depending on the property, adults are willing to use social category information for determining whether two individuals will share both types of novel properties unlike past work had suggested.

**Conclusions.** Taken together, the results of these experiments unequivocally demonstrate that adults exhibit hallmarks of the essentialist bias in their reasoning about social categories. Again, as outlined in the discussion of Aim 1, this forces us to reconsider the developmental trajectory of essentialism. It is increasingly obvious that essentialism does not truly diminish across development, but instead simply becomes a bit harder to detect. It is also increasingly clear that Haslam and colleague’s dimension of social essentialism may be highly valuable for making predictions about essentialist reasoning for a range of social categories. Although we have found strong evidence for this with four particular categories, it is certainly possible that this may apply to a wide swath of social groups. Further, while we specifically chose two groups of categories which rated highly on one dimension and not on another, it is unclear how “super-essential” categories which rate highly on both naturalness and cohesiveness would pattern in these tasks. It may be that such categories would simply show all three forms of essentialist reasoning examined here with no selectivity, but further investigation will be necessary to substantiate such a claim.

**Objections to the Essentialist Account**
Although we have proposed and demonstrated through various tasks that adults do exhibit essentialist reasoning about the specific nature of social groups, there is not universal consensus around the essentialist proposal. Notably, Strevens (2000) proposes the “minimal hypothesis” as an alternative interpretation for the evidence that has been presented in favor of essentialism. The minimal hypothesis states that especially children represent natural kinds and their properties as being connected by kind-laws (K-laws) which are causal in nature. That is, we need only view kind membership as being the cause of properties in order to represent those kinds. Strevens argues that, in effect, this bypasses the need to represent an essence at all, but rather the representation is the causal link between category membership and observed properties.

Such K-laws could arguably account for aspects of essentialist reasoning like inductive potential and even inherent potential. For instance, a causal rule like “if an individual is a girl, then they will want to play with tea sets,” could be general enough to account for inferences from one girl to another in the case of inductive potential and be temporally vague enough to account for the developmentally delayed nature of inherent potential (i.e., “an individual who is a girl will eventually want to play with tea sets”). If the types of responses seen in these tasks could be characterized as simply the result of the outputs of various causal chains, then we would have to reconsider the notion that individuals hold specific assumptions about the nature of categories.

However, this view requires many K-laws to represent all possible category-based properties. Further, the dual-processing evidence presented here seems to generally discount this notion. It is unclear how or why placing participants under time pressure would change the fundamental nature of K-laws to produce different responses. Although one could argue that adults have two sets of K-laws – one that mirrors the promiscuous, category-based responses
generated by children, and one that represents the more selective category-based responding seen by adults at baseline – this creates a problem of combinatorial explosion. The number of K-laws necessary to represent all properties of all categories – especially one set that arises early in development and is retained into adulthood and a second, selective set that arises in adulthood – would be enormous. A single, ever-present root cause such as an essence which may be suppressed by explicit reasoning processes would seem much a more plausible and parsimonious explanation for these data.

Lastly, it is not clear how such unidirectional causal links might account for aspects of essentialism like resistance to change. Once the link between category and properties have been severed through some type of change, it is no longer clear how category membership may be represented. Indeed, in response to Strevens (2000) Ahn et al (2001) argue that the minimal hypothesis cannot account for essentialist biases in relation to non-natural causal laws (e.g. a tiger who got sick and was put in a lion costume, but children still view as a tiger – Keil, 1989) nor does it account for a seemingly coherent account of categories that might arise from a singular cause (an essence) as opposed to a collection of K-laws. Lastly, Ahn et al propose that the K-laws of the minimal hypothesis could in and of themselves be conceptualized as an essence, especially if the K-law was simply “something about being a tiger causes it to have stripes.”

Here, the level of representation is important because Strevens seems to privilege causal laws in the sense that the existence of them precludes the possibility of essence. Conversely, Ahn et al’s view could be characterized as causal laws (or at least the causal nature of an essence) can be represented as part of the essence itself. Indeed, the mere suggestion of a causal mechanism for the expression of properties is perfectly in line with the essentialist account of categorization.
We need not know what it is about being a male or female that gives rise to certain behaviors or physical properties (though we may have an intuitive idea of what such an essence might be), but the belief that some underlying causal force exists at all is consistent with representation of an essence.

**Future Directions**

The experiments outlined here have raised three particular questions for future study. The first is whether the evidence for the dual processing account of essentialism can be found developmentally. While the study of adult participants provided the ideal testing ground for determining whether essentialism would re-emerge under certain conditions (Eidson & Coley, 2014), it is both plausible and expected that similar results should be seen among older children and teenagers.

As both Taylor and colleagues (2009) and Rhodes and Gelman (2008) have shown, essentialist responding tends to decrease gradually over the course of development. Young children around age five or six are relatively promiscuous essentialists, but this begins to decline by age eight to ten. By the teenage years, essentialist responding has decreased markedly until adulthood shows little or selective essentialist responding. However, given that we now know that essentialist responding is ready and waiting to re-emerge in adults, we would expect teens, for instance, should also exhibit increases in essentialist responding under time pressure to closer resemble their younger counterparts. It seems likely that such increases should be observable at all age groups with speeded essentialist responding always closely representing the next youngest age group. However, the effect would eventually hit a ceiling with young children around age 5 to 6 who already exhibit strong, promiscuous essentialist responding.
It is also an open question what older adult participants do at baseline or would do under time pressure in various essentialist tasks. Much as Eidson and Coley showed that development of essentialist reasoning does not simply arbitrarily stop when individuals reach adulthood, it seems unlikely that changes in reasoning would not be viewed across the lifespan. Thus, if developmental trends like those observed by Taylor et al continued, it could be that essentialist responding completely bottoms out in older adults. However, we would still expect that when strained via time pressure or cognitive load, older adults would still exhibit greater essentialist responding.

The second question that is important to examine in future work concerns the selective essentialist responding observed in Experiments 4 and 5 for resistance to transformation and inductive potential, respectively. These experiments were a first attempt to establish the baseline essentialist responding for adults in such tasks. Although they certainly reveal adults’ essentialist tendencies, they reveal selective response patterns not unlike those seen in past work on switched-at-birth tasks (Taylor 1996; Taylor et al, 2009). Although the patterns viewed in Experiments 4 and 5 were predicted by Haslam and colleagues’ (2000) dimensions of essentialism which may suggest that they are in fact the true extent of adults’ essentialist reasoning, it is also possible that time pressure or cognitive load manipulations could affect responding for these tasks just as they do for the switched-at-birth task. If this were the case, we might predict that under time pressure, categories would become overall more resistant to change and have more inductive potential. Whether such increases would wipe out existing interactions between transformation conditions or properties and the category being examined is an open empirical question.
The third question, also related to Experiment 4 and 5, is how children might respond to the social transformation task or dyadic inference task. Past work has suggested that children hold essentialist beliefs about natural kinds’ resistance to transformation (e.g., Keil, 1989) but Experiment 4 was the first to examine beliefs about social transformations. Given that we saw selective patterns of essentialism for adults in the transformation task, just as adults have previously exhibited selective patterns of essentialism in a SAB task while children were promiscuous essentialists (e.g., Taylor et al, 2009), it may be that children hold promiscuous essentialist beliefs about resistance to change as well ignoring the naturalness and cohesiveness distinctions. The same may also be true of children’s reasoning in the dyadic inference task where they may exhibit inductive potential for all categories and equally for both behaviors and physical properties. Again, this is an empirical question.

**Conclusion**

The current work has shown that adults indeed exhibit social essentialist tendencies. They exhibit multiple aspects of essentialist reasoning for various social categories. Further, this essentialist reasoning can often be predicted by various individual differences among adults. Thus, we must reconsider the notion that social essentialism dissipates into adulthood in order to better understand how our essentialist tendencies influence the ways in think about and interact with our social world.
References


Taylor, M. G., Rhodes, M., & Gelman, S. A. (2009). Boys will be boys; Cows will be cows: Children’s essentialist reasoning about gender categories and animal species. Child Development, 80(2), 461-481.


Appendix A: Switched-at-Birth Vignettes for Experiment 3

Race

Black Child Raised by White Parents. Two families (one who was white and one who was black) gave birth to infants in the same hospital on the same day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the white family brought home the child born to the black family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.

Please think about the child born to the black family but raised by the white family. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the black family gave birth to the child, but the white family raised the child.

White Child Raised by Black Parents. Two families (one who was white and one who was black) gave birth to infants in the same hospital on the same day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the black family brought home the child born to the white family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.

Please think about the child born to the white family but raised by the black family. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the white family gave birth to the child, but the black family raised the child.

Black Child Raised by Hispanic Parents. Two families (one who was black and one who was hispanic) gave birth to infants in the same hospital on the same day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the hispanic family brought home the child born to the black family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.

Please think about the child born to the black family but raised by the hispanic family. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the black family gave birth to the child, but the hispanic family raised the child.

Hispanic Child Raised by Black Parents. Two families (one who was black and one who was hispanic) gave birth to infants in the same hospital on the same day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the black family brought home the child born to the hispanic family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.
Please think about the child **born to the hispanic family but raised by the black family**. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the hispanic family **gave birth** to the child, but the black family **raised** the child.

**Hispanic Child Raised by White Parents.** Two families (one who was white and one who was hispanic) gave birth to infants in the same hospital on the same day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the white family brought home the child born to the hispanic family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.

Please think about the child **born to the hispanic family but raised by the white family**. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the hispanic family **gave birth** to the child, but the white family **raised** the child.

**White Child Raised by Hispanic Parents.** Two families (one who was white and one who was hispanic) gave birth to infants in the same hospital on the same day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the hispanic family brought home the child born to the white family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.

Please think about the child **born to the white family but raised by the hispanic family**. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the white family **gave birth** to the child, but the hispanic family **raised** the child.

**Gender**

**Female Child Raised by Male Family.** Two families (one with a single father and several sons and one with a single mother and several daughters) adopted an infant in the same hospital on the same day. The family with a single father and several sons adopted a healthy baby girl while the family with a single mother and several daughters adopted a healthy baby boy. Both children grew up happy and well loved.

Please think about the **infant who was a girl** but **raised by the family with a single father and several sons**. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the **infant** was a girl, but the family with the single father and several sons **raised** the child.
Male Child Raised by Female Family. Two families (one with a single father and several sons and one with a single mother and several daughters) adopted an infant in the same hospital on the same day. The family with a single father and several sons adopted a healthy baby girl while the family with a single mother and several daughters adopted a healthy baby boy. Both children grew up happy and well loved.

Please think about the infant who was a boy but raised by the family with a single mother and several daughters. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the infant was a boy, but the family with the single mother and several daughters raised the child.

Religion

Jewish Child Raised by Christian Parents. Two families (one who was Christian and one who was Jewish) gave birth to infants in the same hospital on the same day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the Christian family brought home the child born to the Jewish family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.

Please think about the child born to the Jewish family but raised by the Christian family. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the Jewish family gave birth to the child, but the Christian family raised the child.

Christian Child Raised by Jewish Parents. Two families (one who was Christian and one who was Jewish) gave birth to infants in the same hospital on the same day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the Jewish family brought home the child born to the Christian family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.

Please think about the child born to the Christian family but raised by the Jewish family. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the Christian family gave birth to the child, but the Jewish family raised the child.

Jewish Child Raised by Muslim Parents. Two families (one who was Muslim and one who was Jewish) gave birth to infants in the same hospital on the same day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the Muslim family brought home the child born to the Jewish family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.
ADULTS’ SOCIAL ESSENTIALISM

Please think about the child born to the Jewish family but raised by the Muslim family. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the Jewish family gave birth to the child, but the Muslim family raised the child.

**Muslim Child Raised by Jewish Parents.** Two families (one who was Muslim and one who was Jewish) gave birth to infants in the same hospital on the same day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the Jewish family brought home the child born to the Muslim family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.

Please think about the child born to the Muslim family but raised by the Jewish family. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the Muslim family gave birth to the child, but the Jewish family raised the child.

**Muslim Child Raised by Christian Parents.** Two families (one who was Christian and one who was Muslim) gave birth to infants in the same hospital on the same day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the Christian family brought home the child born to the Muslim family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.

Please think about the child born to the Muslim family but raised by the Christian family. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the Muslim family gave birth to the child, but the Christian family raised the child.

**Christian Child Raised by Muslim Parents.** Two families (one who was Christian and one who was Muslim) gave birth to infants in the same hospital on the same day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the Muslim family brought home the child born to the Christian family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.

Please think about the child born to the Christian family but raised by the Muslim family. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the Christian family gave birth to the child, but the Muslim family raised the child.

**Political Affiliation**

**Republican Child Raised by Democrat Parents.** Two families (one who were Democrats and one who were Republicans) gave birth to infants in the same hospital on the same
day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the Democrat family brought home the child born to the Republican family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.

Please think about the child **born to the Republican family but raised by the Democrat family**. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the Republican family *gave birth* to the child, but the Democrat family *raised* the child.

**Democrat Child Raised by Republican Parents.** Two families (one who were Democrats and one who were Republicans) gave birth to infants in the same hospital on the same day. Without anyone knowing, the children were mistakenly switched before they were taken home, so that the Republican family brought home the child born to the Democrat family and vice versa. The mistake was never discovered, and both children grew up happy and well loved.

Please think about the child **born to the Democrat family but raised by the Republican family**. Use this information to make your best guess about what characteristics this child would have at your age. Choose the answer that seems right to you.

Remember, the Democrat family *gave birth* to the child, but the Republican family *raised* the child.
Appendix B: Transformation Vignettes for Experiment 4

Race

**Stereotypical Behaviors.** Olivia was a healthy black person. She liked to listen to rap music, was good at basketball, and was part of a hip-hop dance group. One day, Olivia decided to change to be like a white person instead. She started to listen to country music, gave up basketball for golf, and began to take ballet dance classes every week.

**Physical.** Olivia was a healthy black person. She liked to listen to rap music, was good at basketball, and was part of a hip-hop dance group. One day, Olivia decided to change to be like a white person instead. She took medication that lightened her skin, had her hair permanently straightened, and had plastic surgery to make her nose and lips narrower. She continued to listen to rap music, was still a successful basketball player, and continued be part of a hip-hop dance group.

**Both.** Olivia was a healthy black person. She liked to listen to rap music, was good at basketball, and was part of a hip-hop dance group. One day, Olivia decided to change to be like a white person instead. She took medication that lightened her skin, had her hair permanently straightened, and had plastic surgery to make her nose and lips narrower. She started to listen to country music, gave up basketball for golf, and began to take ballet dance classes every week.

Gender

**Stereotypical Behaviors.** Mike was a healthy young man. He liked to wear shirts and ties, watch football, and drink beer. One day he decided to change to be like a young woman instead. He asked to be called “Michelle,” started wearing dresses, watching “Grey’s Anatomy,” and drinking white wine.
Physical. Mike was a healthy young man. He liked to wear shirts and ties, watch football, and drink beer. One day he decided to change to be like a young woman instead. He had his genitals surgically transformed into those of a young woman and started taking hormone treatments. Mike continued to wear shirts and ties, watch football, and drink beer.

Both. Mike was a healthy young man. He liked to wear shirts and ties, watch football, and drink beer. One day he decided to change to be like a young woman instead. He had his genitals surgically transformed into those of a young woman and started taking hormone treatments. He asked to be called “Michelle,” started wearing dresses, watching “Grey’s Anatomy,” and drinking white wine.

Religion

Characteristic Behaviors. Lisa was a healthy Muslim. She attended services at a Mosque, prayed five times a day, and observed holidays such as Ramadan. One day, she decided to change to be like a Christian instead. She began to attend church every Sunday, met with a Bible study group, and observed holidays such as Christmas and Easter.

Nominal Properties. Lisa was a healthy Muslim. She attended services at a Mosque, prayed five times a day, and observed holidays such as Ramadan. One day, she decided to change to be like a Christian instead. She went through Christian confirmation, was baptized, and took her first Communion. She still attends services at a Mosque, prays five times a day, and observes holidays such as Ramadan.

Both. Lisa was a healthy Muslim. She attended services at a Mosque, prayed five times a day, and observed holidays such as Ramadan. One day, she decided to change to be like a Christian instead. She went through Christian confirmation, was baptized, and took her first
Communion. She began to attend church every Sunday, met with a Bible study group, and observed holidays such as Christmas and Easter.

**Political Affiliation**

**Characteristic Behaviors.** John was a healthy Democrat. He attended pro-choice rallies, volunteered for a local Democratic candidate’s campaign, and lived in San Francisco. One day, John decided to change to be like a Republican instead. He began to attend pro-life rallies, volunteered for a local Republican candidate’s campaign, and moved to a small town in Alabama.

**Nominal Properties.** John was a healthy Democrat. He attended pro-choice rallies, volunteered for a local Democratic candidate’s campaign, and lived in San Francisco. One day, John decided to change to be like a Republican instead. He began to adopt conservative values and had his voter identification changed to “registered Republican.” He still attends pro-choice rallies, volunteers for a local Democratic candidate’s campaign, and lives in San Francisco.

**Both.** John was a healthy Democrat. He attended pro-choice rallies, volunteered for a local Democratic candidate’s campaign, and lived in San Francisco. One day, John decided to change to be like a Republican instead. He began to adopt conservative values and had his voter identification changed to “registered Republican.” He began to attend pro-life rallies, volunteered for a local Republican candidate’s campaign, and moved to a small town in Alabama.
Appendix C: Target Figures for Experiment 5

Black Female Figures With Political (Democrat / Republican) and Religious (Christian / Muslim) Iconography.

Black Male Figures With Political (Democrat / Republican) and Religious (Christian / Muslim) Iconography.
White Female Figures With Political (Democrat / Republican) and Religious (Christian / Muslim) Iconography.

White Male Figures With Political (Democrat / Republican) and Religious (Christian / Muslim) Iconography.
<table>
<thead>
<tr>
<th>Property Type</th>
<th>Male Property</th>
<th>Female Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioral</td>
<td>Play with a toy truck</td>
<td>Play with a tea set</td>
</tr>
<tr>
<td></td>
<td>Play with baseball cards</td>
<td>Play dress-up</td>
</tr>
<tr>
<td></td>
<td>Collect tools and nails</td>
<td>Collect dolls</td>
</tr>
<tr>
<td>Preference</td>
<td>Will be a construction worker</td>
<td>Will be a nursery school teacher</td>
</tr>
<tr>
<td></td>
<td>Like to build things</td>
<td>Like to sew</td>
</tr>
<tr>
<td></td>
<td>Like to go fishing</td>
<td>Like to put on make-up</td>
</tr>
<tr>
<td></td>
<td>Want to be a firefighter</td>
<td>Want to be a nurse</td>
</tr>
<tr>
<td></td>
<td>Want to be a football player</td>
<td>Want to be a ballet dancer</td>
</tr>
<tr>
<td>Physical</td>
<td>Have boy blood inside</td>
<td>Have girl blood inside</td>
</tr>
<tr>
<td></td>
<td>Have a boy’s body</td>
<td>Have a girl’s body</td>
</tr>
<tr>
<td></td>
<td>Have a boy’s brain</td>
<td>Have a girl’s brain</td>
</tr>
<tr>
<td></td>
<td>Have a boy’s heart</td>
<td>Have a girl’s heart</td>
</tr>
<tr>
<td>Category-based</td>
<td>Grow up to be a daddy</td>
<td>Grow up to be a mommy</td>
</tr>
<tr>
<td></td>
<td>Grow up to be a husband</td>
<td>Grow up to be a wife</td>
</tr>
<tr>
<td></td>
<td>Grow up to have a beard</td>
<td>Grow up to have breasts</td>
</tr>
<tr>
<td></td>
<td>Have a deep and low voice</td>
<td>Have the same voice</td>
</tr>
</tbody>
</table>

Table 1. Property pairs used in the Switched-At-Birth task for Experiments 1 and 2.
<table>
<thead>
<tr>
<th>Individual Difference</th>
<th>Variable Predicted</th>
<th>Direction of Association</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variability - Behavior</td>
<td>Speeded Responding</td>
<td>+</td>
</tr>
<tr>
<td>Variability - Physical</td>
<td>Speeded Responding</td>
<td>+</td>
</tr>
<tr>
<td>Motivation - Internal</td>
<td>Degree of Change</td>
<td>+</td>
</tr>
<tr>
<td>Motivation - External</td>
<td>Degree of Change</td>
<td>+</td>
</tr>
<tr>
<td>Political Affiliation</td>
<td>Speeded Responding</td>
<td>+</td>
</tr>
<tr>
<td>RWA</td>
<td>Speeded Responding</td>
<td>+</td>
</tr>
<tr>
<td>Masculinity</td>
<td>Speeded Responding</td>
<td>+</td>
</tr>
<tr>
<td>Femininity</td>
<td>Speeded Responding</td>
<td>+</td>
</tr>
<tr>
<td>Religiosity</td>
<td>Speeded Responding</td>
<td>+</td>
</tr>
<tr>
<td>Stroop Error Difference</td>
<td>Degree of Change</td>
<td>-</td>
</tr>
<tr>
<td>Stroop RT Difference</td>
<td>Degree of Change</td>
<td>-</td>
</tr>
<tr>
<td>SIC Error Difference</td>
<td>Degree of Change</td>
<td>-</td>
</tr>
<tr>
<td>SIC RT Difference</td>
<td>Degree of Change</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2. Predicted directions of association between individual differences and speeded essentialist responding or degree of change in essentialist responding.
Table 3. Beta (β) coefficients for individual difference predictors in Experiment 2. †p < .10, *p < .05, **p < .01, ***p < .001

<table>
<thead>
<tr>
<th>Participants and Variable</th>
<th>Variability Behavior</th>
<th>Variability Physical</th>
<th>Motivation Internal</th>
<th>Motivation External</th>
<th>Political Affiliation</th>
<th>RWA</th>
<th>Masculinity</th>
<th>Femininity</th>
<th>Stroop Error</th>
<th>Stroop RT</th>
<th>SIC Error</th>
<th>SIC RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>All – Degree of Change</td>
<td>-0.118</td>
<td>0.184</td>
<td>-0.070</td>
<td>-0.080</td>
<td>-0.049</td>
<td>0.048</td>
<td>0.06</td>
<td>-0.084</td>
<td>0.065</td>
<td>-0.119</td>
<td>0.076</td>
<td>-0.101</td>
</tr>
<tr>
<td>All – Speeded Responses</td>
<td>-0.199</td>
<td>0.236</td>
<td>-0.076</td>
<td>0.011</td>
<td>-0.105</td>
<td>0.140</td>
<td>0.242†</td>
<td>0.098</td>
<td>0.001</td>
<td>0.110</td>
<td>0.100</td>
<td>-0.046</td>
</tr>
<tr>
<td>Females – Degree of Change</td>
<td>-0.521</td>
<td>0.226</td>
<td>-0.132</td>
<td>-0.247</td>
<td>0.453</td>
<td>-0.284</td>
<td>-0.129</td>
<td>-0.059</td>
<td>0.120</td>
<td>-0.186</td>
<td>0.070</td>
<td>-0.018</td>
</tr>
<tr>
<td>Females – Speeded Responses</td>
<td>-0.368</td>
<td>-0.043</td>
<td>-0.094</td>
<td>-0.278</td>
<td>0.551*</td>
<td>-0.263</td>
<td>0.033</td>
<td>0.121</td>
<td>0.139</td>
<td>-0.094</td>
<td>-</td>
<td>0.211</td>
</tr>
<tr>
<td>Males – Degree of Change</td>
<td>0.331†</td>
<td>-0.133</td>
<td>-0.354†</td>
<td>0.323</td>
<td>-0.517**</td>
<td>0.258</td>
<td>0.349†</td>
<td>0.090</td>
<td>0.266</td>
<td>-0.034</td>
<td>-</td>
<td>-0.348</td>
</tr>
<tr>
<td>Males – Speeded Responses</td>
<td>-0.035</td>
<td>0.173</td>
<td>-0.244</td>
<td>0.167</td>
<td>-0.457**</td>
<td>0.217</td>
<td>0.276</td>
<td>0.114</td>
<td>-0.006</td>
<td>0.015</td>
<td>0.049</td>
<td>0.148</td>
</tr>
</tbody>
</table>
### Table 4. Properties used for the Switched-at-Birth task in Experiment 3.

<table>
<thead>
<tr>
<th>Behavioral</th>
<th>Physical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest in Activities</td>
<td>Organ Types</td>
</tr>
<tr>
<td>Political Beliefs</td>
<td>Blood Type</td>
</tr>
<tr>
<td>Hobbies</td>
<td>Personality Traits</td>
</tr>
<tr>
<td>Preference for Professions</td>
<td>Genes</td>
</tr>
<tr>
<td>World view / Ideological Similarities</td>
<td>Physical Features</td>
</tr>
<tr>
<td>Goals and Beliefs</td>
<td>Physical Abilities</td>
</tr>
<tr>
<td>Category</td>
<td>Participants and Variable</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td>All – Degree of Change</td>
</tr>
<tr>
<td></td>
<td>Female – Degree of Change</td>
</tr>
<tr>
<td></td>
<td>Male – Degree of Change</td>
</tr>
<tr>
<td></td>
<td>All – Speeded Responses</td>
</tr>
<tr>
<td></td>
<td>Female – Speeded Responses</td>
</tr>
<tr>
<td></td>
<td>Male – Speeded Responses</td>
</tr>
<tr>
<td>Gender</td>
<td>All – Degree of Change</td>
</tr>
<tr>
<td></td>
<td>Female – Degree of Change</td>
</tr>
<tr>
<td></td>
<td>Male – Degree of Change</td>
</tr>
<tr>
<td></td>
<td>All – Speeded Responses</td>
</tr>
<tr>
<td></td>
<td>Female – Speeded Responses</td>
</tr>
<tr>
<td></td>
<td>Male – Speeded Responses</td>
</tr>
<tr>
<td>Religion</td>
<td>All – Degree of Change</td>
</tr>
<tr>
<td></td>
<td>Female – Degree of Change</td>
</tr>
<tr>
<td></td>
<td>Male – Degree of Change</td>
</tr>
<tr>
<td></td>
<td>All – Speeded Responses</td>
</tr>
<tr>
<td></td>
<td>Female – Speeded Responses</td>
</tr>
<tr>
<td></td>
<td>Male – Speeded Responses</td>
</tr>
<tr>
<td>Political Affiliation</td>
<td>All – Degree of Change</td>
</tr>
<tr>
<td></td>
<td>Female – Degree of Change</td>
</tr>
<tr>
<td></td>
<td>Male – Degree of Change</td>
</tr>
<tr>
<td></td>
<td>All – Speeded Responses</td>
</tr>
<tr>
<td></td>
<td>Female – Speeded Responses</td>
</tr>
<tr>
<td></td>
<td>Male – Speeded Responses</td>
</tr>
</tbody>
</table>

Table 5. Beta (β) coefficients for individual difference predictors in Experiment 3. † p < .10, * p < .05, ** p < .01, *** p < .001
<table>
<thead>
<tr>
<th>Category</th>
<th>Stereotypical / Characteristic Behaviors</th>
<th>Physical / Nominal Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race (Black – White)</td>
<td>Attire, taste in TV show, favorite beverage</td>
<td>Genitals, hormones</td>
</tr>
<tr>
<td>Gender (Male – Female)</td>
<td>Taste in music, taste in sports, type of dancing</td>
<td>Skin color, hair type, facial features</td>
</tr>
<tr>
<td>Religion (Muslim – Christian)</td>
<td>Type of religious service, holidays, Bible study instead of praying five times a day</td>
<td>Went through Confirmation, got baptized, took first communion</td>
</tr>
<tr>
<td>Political Affiliation (Democrat – Republican)</td>
<td>Type of rallies attended, party of candidate campaigned for, city lived in</td>
<td>Change in from liberal to conservative values, voter identification</td>
</tr>
</tbody>
</table>

Table 6. Summary of properties used for each category and transformation condition in Experiment 4.
Table 7. Beta (β) coefficients for individual difference predictors for race and gender in Experiment 4. †p < .10, *p < .05, **p < .01, ***p < .001
<table>
<thead>
<tr>
<th>Category</th>
<th>Participants and Condition</th>
<th>Motivation – Internal</th>
<th>Motivation – External</th>
<th>Political Affiliation</th>
<th>RWA</th>
<th>Masculinity</th>
<th>Femininity</th>
<th>Religiosity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Religion All Participants &amp; Conditions</td>
<td>0.105</td>
<td>-0.142</td>
<td>-0.047</td>
<td>0.079</td>
<td>0.051</td>
<td>0.028</td>
<td>-0.104</td>
<td></td>
</tr>
<tr>
<td>Females (All Conditions)</td>
<td>0.178</td>
<td>-0.176</td>
<td>-0.277</td>
<td>0.252</td>
<td>-0.143</td>
<td>0.063</td>
<td>0.088</td>
<td></td>
</tr>
<tr>
<td>Males (All Conditions)</td>
<td>0.137</td>
<td>-0.105</td>
<td>0.060</td>
<td>0.056</td>
<td>0.179</td>
<td>-0.021</td>
<td>-0.273</td>
<td></td>
</tr>
<tr>
<td>Characteristic (Females &amp; Males)</td>
<td>0.085</td>
<td>-0.158</td>
<td>-0.056</td>
<td>-0.135</td>
<td>0.038</td>
<td>-0.031</td>
<td>0.160</td>
<td></td>
</tr>
<tr>
<td>Nominal (Females &amp; Males)</td>
<td>0.139</td>
<td>0.064</td>
<td>0.202</td>
<td>-0.223</td>
<td>-0.005</td>
<td>0.039</td>
<td>-0.083</td>
<td></td>
</tr>
<tr>
<td>Both (Females &amp; Males)</td>
<td>0.247</td>
<td>-0.271</td>
<td>0.073</td>
<td>0.320</td>
<td>-0.071</td>
<td>0.108</td>
<td>-0.347</td>
<td></td>
</tr>
<tr>
<td>Female Characteristic</td>
<td>0.346</td>
<td>0.346</td>
<td>-0.714</td>
<td>0.597</td>
<td>-0.525</td>
<td>-0.320</td>
<td>0.433</td>
<td></td>
</tr>
<tr>
<td>Female Nominal</td>
<td>0.060</td>
<td>0.236</td>
<td>-0.253</td>
<td>-0.051</td>
<td>-0.033</td>
<td>0.488</td>
<td>-0.197</td>
<td></td>
</tr>
<tr>
<td>Female Both</td>
<td>0.151</td>
<td>-0.387</td>
<td>-0.214</td>
<td>0.384</td>
<td>-0.110</td>
<td>0.089</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>Male Characteristic</td>
<td>0.093</td>
<td>-0.304</td>
<td>0.228</td>
<td>-0.602</td>
<td>0.357</td>
<td>-0.139</td>
<td>0.322</td>
<td></td>
</tr>
<tr>
<td>Male Nominal</td>
<td>-0.016</td>
<td>-0.030</td>
<td>0.479</td>
<td>-0.679</td>
<td>0.052</td>
<td>-0.165</td>
<td>0.242</td>
<td></td>
</tr>
<tr>
<td>Male Both</td>
<td>0.054</td>
<td>0.002</td>
<td>0.238</td>
<td>0.197</td>
<td>-0.239</td>
<td>0.112</td>
<td>-0.603†</td>
<td></td>
</tr>
<tr>
<td>Political Affiliation All Participants &amp; Conditions</td>
<td>-0.069</td>
<td>0.121</td>
<td>-0.040</td>
<td>-0.033</td>
<td>-0.013</td>
<td>0.063</td>
<td>-0.064</td>
<td></td>
</tr>
<tr>
<td>Females (All Conditions)</td>
<td>0.061</td>
<td>0.021</td>
<td>-0.084</td>
<td>0.079</td>
<td>-0.137</td>
<td>0.208</td>
<td>-0.062</td>
<td></td>
</tr>
<tr>
<td>Males (All Conditions)</td>
<td>-0.144</td>
<td>0.182</td>
<td>-0.101</td>
<td>-0.093</td>
<td>0.149</td>
<td>-0.171</td>
<td>-0.084</td>
<td></td>
</tr>
<tr>
<td>Characteristic (Females &amp; Males)</td>
<td>-0.123</td>
<td>0.102</td>
<td>-0.049</td>
<td>-0.185</td>
<td>0.109</td>
<td>0.218</td>
<td>-0.034</td>
<td></td>
</tr>
<tr>
<td>Nominal (Females &amp; Males)</td>
<td>0.207</td>
<td>0.251</td>
<td>0.253</td>
<td>-0.142</td>
<td>-0.141</td>
<td>-0.067</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td>Both (Females &amp; Males)</td>
<td>-0.152</td>
<td>0.100</td>
<td>0.007</td>
<td>0.057</td>
<td>-0.148</td>
<td>0.101</td>
<td>-0.201</td>
<td></td>
</tr>
<tr>
<td>Female Characteristic</td>
<td>-0.008</td>
<td>-0.008</td>
<td>-0.486</td>
<td>0.418</td>
<td>0.058</td>
<td>0.017</td>
<td>-0.073</td>
<td></td>
</tr>
<tr>
<td>Female Nominal</td>
<td>0.223</td>
<td>0.407†</td>
<td>0.053</td>
<td>0.061</td>
<td>-0.267</td>
<td>0.533†</td>
<td>-0.169</td>
<td></td>
</tr>
<tr>
<td>Female Both</td>
<td>-0.061</td>
<td>-0.213</td>
<td>-0.108</td>
<td>0.129</td>
<td>-0.242</td>
<td>0.188</td>
<td>-0.043</td>
<td></td>
</tr>
<tr>
<td>Male Characteristic</td>
<td>-0.314</td>
<td>-0.188</td>
<td>-0.020</td>
<td>-0.701</td>
<td>0.232</td>
<td>-0.131</td>
<td>0.507</td>
<td></td>
</tr>
<tr>
<td>Male Nominal</td>
<td>-0.070</td>
<td>0.241</td>
<td>0.145</td>
<td>-0.685</td>
<td>0.092</td>
<td>-0.443</td>
<td>0.452</td>
<td></td>
</tr>
<tr>
<td>Male Both</td>
<td>-0.334</td>
<td>0.541</td>
<td>-0.023</td>
<td>0.159</td>
<td>-0.041</td>
<td>-0.179</td>
<td>-0.450</td>
<td></td>
</tr>
</tbody>
</table>

Table 8. Beta (β) coefficients for individual difference predictors for religion and political affiliation in Experiment 4. †p < .10, *p < .05, **p < .01, ***p < .001
<table>
<thead>
<tr>
<th>Category</th>
<th>Behavior</th>
<th></th>
<th>Physical</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Match</td>
<td>Mismatch</td>
<td>Match</td>
<td>Mismatch</td>
</tr>
<tr>
<td>Race</td>
<td>5.69</td>
<td>4.83</td>
<td>6.17</td>
<td>4.27</td>
</tr>
<tr>
<td>Gender</td>
<td>5.60</td>
<td>4.93</td>
<td>5.67</td>
<td>4.77</td>
</tr>
<tr>
<td>Religion</td>
<td>6.20</td>
<td>4.33</td>
<td>5.66</td>
<td>4.77</td>
</tr>
<tr>
<td>Political Affiliation</td>
<td>6.13</td>
<td>4.40</td>
<td>5.58</td>
<td>4.85</td>
</tr>
</tbody>
</table>

Table 9. Mean likelihood ratings for each social category by property and match conditions in Experiment 5.
<table>
<thead>
<tr>
<th>Category</th>
<th>Participants and Property</th>
<th>Individual Difference Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Motivation – Internal</td>
</tr>
<tr>
<td>Race</td>
<td>All – Behavior</td>
<td>-0.207*</td>
</tr>
<tr>
<td></td>
<td>Female – Behavior</td>
<td>-0.316*</td>
</tr>
<tr>
<td></td>
<td>Male – Behavior</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>All – Physical</td>
<td>-0.041</td>
</tr>
<tr>
<td></td>
<td>Female – Physical</td>
<td>-0.212</td>
</tr>
<tr>
<td></td>
<td>Male – Physical</td>
<td>0.056</td>
</tr>
<tr>
<td>Gender</td>
<td>All – Behavior</td>
<td>-0.085</td>
</tr>
<tr>
<td></td>
<td>Female – Behavior</td>
<td>-0.301*</td>
</tr>
<tr>
<td></td>
<td>Male – Behavior</td>
<td>0.102</td>
</tr>
<tr>
<td></td>
<td>All – Physical</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>Female – Physical</td>
<td>-0.258†</td>
</tr>
<tr>
<td></td>
<td>Male – Physical</td>
<td>0.184</td>
</tr>
<tr>
<td>Religion</td>
<td>All – Behavior</td>
<td>-0.061</td>
</tr>
<tr>
<td></td>
<td>Female – Behavior</td>
<td>-0.260†</td>
</tr>
<tr>
<td></td>
<td>Male – Behavior</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>All – Physical</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td>Female – Physical</td>
<td>-0.016</td>
</tr>
<tr>
<td></td>
<td>Male – Physical</td>
<td>0.101</td>
</tr>
<tr>
<td>Political Affiliation</td>
<td>All – Behavior</td>
<td>-0.044</td>
</tr>
<tr>
<td></td>
<td>Female – Behavior</td>
<td>-0.165</td>
</tr>
<tr>
<td></td>
<td>Male – Behavior</td>
<td>-0.039</td>
</tr>
<tr>
<td></td>
<td>All – Physical</td>
<td>-0.102</td>
</tr>
<tr>
<td></td>
<td>Female – Physical</td>
<td>-0.232</td>
</tr>
<tr>
<td></td>
<td>Male – Physical</td>
<td>0.020</td>
</tr>
</tbody>
</table>

Table 10. Beta (β) coefficients for individual difference predictors in Experiment 5. † p < .10, * p < .05, ** p < .01, *** p < .001
Figure 1. Mean essentialist responses by property type and loading condition for Experiment 1.
Figure 2. Mean reaction time by property and response type for Experiment 1.
Figure 3. Mean essentialist responses by property type and timing condition for Experiment 2.
Figure 4. Counts of individual response patterns in each timing condition for Experiment 2.
Figure 5. Proportion of essentialist responses for each category and property type for Experiment 3.
Figure 6. Counts for individual response patterns in each timing condition for Experiment 3.
Figure 7. Differences in category judgments by transformation condition for race and gender in Experiment 4.
Figure 8. Differences in category judgments by transformation condition for religion and political affiliation in Experiment 4.
Figure 9. Participant response patterns by transformation condition for the category of race for Experiment 4.
Figure 10. Participant response patterns by transformation condition for the category of *gender* for Experiment 4.
Figure 11. Participant response patterns by transformation condition for the category of religion for Experiment 4.
Figure 12. Participant response patterns by transformation condition for the category of political affiliation for Experiment 4.
How likely is it that the person on the right has wellant cells like the person on the left?

(Very Unlikely) 1 2 3 4 5 6 7 8 9 (Very Likely)

Figure 13. Example of a dyadic trial with the base figure (black female Christian Democrat) on the left and the target figure (white male Christian Democrat) on the right for Experiment 5.
Figure 14. Difference in match and mismatch likelihood ratings by property and category for Experiment 5.