NEED FOR MYSTERY: DEVELOPMENT AND VALIDATION OF A NEW
CONSTRUCT

by

JASMINE MARIE CAREY

B.A., The Ohio State University, 2005
M.A. The University of British Columbia, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in
THE FACULTY OF GRADUATE STUDIES
(Psychology)
THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

December 2012

© Jasmine Marie Carey, 2012
Abstract

This dissertation introduces a new construct I have labeled Need for Mystery and the scale designed to measure this construct. Need for Mystery (nMyst) is a tendency to maintain cherished beliefs by mystifying them. Instead of the systematic approach entailed by science, such individuals justify their beliefs in ways that are not easily susceptible to proof or disproof and fit with a more intuitive way of understanding the world. In order to study nMyst, I have constructed a questionnaire to measure individual differences in the construct. Included are details of the scale construction and four studies conducted to establish its validity as a measure of nMyst.

Study 1 (N = 326) evaluated the factor structure of the initially generated item set. It also included a number of other personality scales to begin to assess the nomological network of potentially related constructs. Study 2 (N = 15,630) was collected on Facebook. The final dataset was large enough to allow for independent tests of a number of CFA models. Included was a direct comparison of Western-heritage participants with Chinese-heritage participants. I also tested the nMyst scale items using an Item Response Theory (IRT) analysis and examined the influence of nMyst on religious conversion. Study 3 (N = 360) extended the nomological net to include measures of political attitudes. Participants were recruited using Mechanical Turk. Study 4 (N = 100) specifically addressed issues of discriminant validity and evaluated the temporal stability of the nMyst scale with a 3-month test-retest reliability. These four studies indicate that nMyst is reliable and valid measure that makes a unique contribution to the literature on individual differences in belief.
Preface

No part of this research has been submitted for publication. Data collected at UBC or through Mechanical Turk was approved by the UBC Human Ethics Research Ethics Board under certificates H97-80141 and H11-01944. Data collected on Facebook was approved by the University of Virginia IRB for the Social and Behavioral Sciences under Dr. Jonathan Haidt, certificate number 2007-0147.
# Table of Contents

Abstract............................................................................................................................... ii
Preface............................................................................................................................... iii
Table of Contents.............................................................................................................. iv
List of Tables .................................................................................................................... v
List of Figures ................................................................................................................... vi
List of Abbreviations ...................................................................................................... vii
Acknowledgments ......................................................................................................... viii
Dedication ....................................................................................................................... ix
Chapter 1: Need for Mystery .......................................................................................... 1  
  Measuring nMyst: What should be included? ................................................................. 5  
  Overview of the Four Studies ....................................................................................... 12  
  Summary ...................................................................................................................... 15
Chapter 2: Item Development ..................................................................................... 17  
  Method ......................................................................................................................... 19  
  Results ......................................................................................................................... 19  
  Discussion ................................................................................................................... 22
Chapter 3: Psychometrics, Demographics, and Religious Conversion ....................... 24  
  Religious Conversion .................................................................................................... 26  
  Method ......................................................................................................................... 27  
  Results ......................................................................................................................... 27  
  Discussion ................................................................................................................... 38
Chapter 4: Political Attitudes ....................................................................................... 44  
  Method ......................................................................................................................... 47  
  Measures ..................................................................................................................... 48  
  Results ......................................................................................................................... 48  
  Discussion ................................................................................................................... 50
Chapter 5: Discriminant Validity and Temporal Stability .......................................... 53  
  Method ......................................................................................................................... 55  
  Measures ..................................................................................................................... 56  
  Results ......................................................................................................................... 56  
  Discussion ................................................................................................................... 58
Chapter 6: General Discussion and Conclusion .......................................................... 60  
  Summary ...................................................................................................................... 66  
  Future Directions and Predictions .............................................................................. 70
Tables ............................................................................................................................. 80
Figures ............................................................................................................................ 87
References ....................................................................................................................... 94
Appendix A ...................................................................................................................... 112
Appendix B ...................................................................................................................... 114
Appendix C ...................................................................................................................... 115
Appendix D ...................................................................................................................... 117
SPQ Items ....................................................................................................................... 117
Mysticism Items ............................................................................................................. 118
List of Tables

Table 1  Factor loadings of the initial 32 item set in Study 1................................. 80
Table 2  Final item set loadings, Study 1................................................................. 81
Table 3  CFA Model fit comparisons from Study 2.................................................... 82
Table 4  IRT discrimination parameters from the two dimensional model and difficulty parameters from the unidimensional models................................. 83
Table 5  IRT parameters for the Chinese heritage sample and DIF measures in comparison to the European-American sample................................................. 84
Table 6  Number of participants in each current religious affiliation by childhood religious affiliation............................................................. 85
Table 7  Internal consistency reliability across ethnic groups, Study 2......................... 86
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Bidimensional representation of nMyst subscales as quadrants</td>
<td>87</td>
</tr>
<tr>
<td>Figure 2</td>
<td>CFA Final Model with standardized estimates, Study 2</td>
<td>88</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Comparison of total information curves for European Heritage and Chinese Heritage samples, Study 2</td>
<td>89</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Means of Anti-Science comparing non-converts to converts to and from each religion</td>
<td>90</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Means of Pro-Magic comparing non-converts to converts to and from each religion</td>
<td>91</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Ethnicity ANOVA results Study 2</td>
<td>92</td>
</tr>
<tr>
<td>Figure 7</td>
<td>ANOVA results of education level by Anti-Science and Pro-Magic, Study 4.</td>
<td>93</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike Information Criterion</td>
<td></td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
<td></td>
</tr>
<tr>
<td>AS</td>
<td>Anti-Science</td>
<td></td>
</tr>
<tr>
<td>BAEM</td>
<td>Bock-Aitken EM Estimation Method</td>
<td></td>
</tr>
<tr>
<td>BFI</td>
<td>Big Five Inventory</td>
<td></td>
</tr>
<tr>
<td>CFA</td>
<td>Confirmatory Factor Analysis</td>
<td></td>
</tr>
<tr>
<td>CFI</td>
<td>Comparative Fit Index</td>
<td></td>
</tr>
<tr>
<td>DIF</td>
<td>Differential Item Functioning</td>
<td></td>
</tr>
<tr>
<td>DSM</td>
<td>Diagnostic and Statistical Manual</td>
<td></td>
</tr>
<tr>
<td>EFA</td>
<td>Exploratory Factor Analysis</td>
<td></td>
</tr>
<tr>
<td>HSP</td>
<td>Human Subjects Pool</td>
<td></td>
</tr>
<tr>
<td>IRT</td>
<td>Item Response Theory</td>
<td></td>
</tr>
<tr>
<td>JWB</td>
<td>Just World Belief</td>
<td></td>
</tr>
<tr>
<td>MAP</td>
<td>Minimum Average Partial</td>
<td></td>
</tr>
<tr>
<td>MFQ</td>
<td>Moral Foundations Questionnaire</td>
<td></td>
</tr>
<tr>
<td>MGCFA</td>
<td>Multi-group Confirmatory Factor Analysis</td>
<td></td>
</tr>
<tr>
<td>MHRM</td>
<td>Metropolis-Hastings Robbins-Monro Monte Carlo Estimation Method</td>
<td></td>
</tr>
<tr>
<td>ML</td>
<td>Maximum Likelihood</td>
<td></td>
</tr>
<tr>
<td>MMM</td>
<td>Meaning Maintenance Model</td>
<td></td>
</tr>
<tr>
<td>nMyst</td>
<td>Need for Mystery</td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>Parallel Analysis</td>
<td></td>
</tr>
<tr>
<td>PM</td>
<td>Pro-Magic</td>
<td></td>
</tr>
<tr>
<td>PNS</td>
<td>Personal Need for Structure</td>
<td></td>
</tr>
<tr>
<td>RMSEA</td>
<td>Root Mean Square Error of Approximation</td>
<td></td>
</tr>
<tr>
<td>RWA</td>
<td>Right Wing Authoritarianism</td>
<td></td>
</tr>
<tr>
<td>SEM</td>
<td>Standard Error of Measurement</td>
<td></td>
</tr>
<tr>
<td>SJT</td>
<td>System Justification Theory</td>
<td></td>
</tr>
<tr>
<td>SPD</td>
<td>Schizotypal Personality Disorder</td>
<td></td>
</tr>
<tr>
<td>SPQ</td>
<td>Schizotypal Personality Questionnaire</td>
<td></td>
</tr>
<tr>
<td>TLI</td>
<td>Tucker-Lewis Fit Index</td>
<td></td>
</tr>
<tr>
<td>TMT</td>
<td>Terror Management Theory</td>
<td></td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
<td></td>
</tr>
</tbody>
</table>
Acknowledgments

I’d like to thank the following people for their support on this project, both personal and professional. My advisor Del Paulhus for his support and guidance throughout my graduate school career and his patience with all my crazy ideas. My labmates; Craig Nathanson, Dan Jones, and Erin Buckels for your knowledge, support, and, most importantly, friendship. My co-advisor Victoria Savalei for showing me that statistics can be fun and for her helpful comments through the thesis writing process. Also my thesis committee members; Lawrence Ward for coming up with, and letting me use, the term “Need for Mystery”, and Steve Heine for helping me work through the relationship between nMyst and Meaning Maintenance and for liking the idea.
Dedication

This thesis is dedicated to my wonderful husband. You’re my inspiration and motivation and I could never have done this without you.

Need for Mystery is dedicated to my Mother, who taught me everything I know about magic.
Chapter 1: Need for Mystery

This dissertation will introduce and elaborate on a new personality construct dubbed Need for Mystery (nMyst). The construct is related to how people form and hold cherished beliefs. My introduction provides a definition of nMyst and lays out its theoretical basis. The main body of the thesis will focus on developing a measure of individual differences in nMyst along with support for its reliability, structure, and validity. I will build the nomological network of correlates and discuss applications to various areas of research.

I will argue that Need for Mystery (nMyst) is a personality trait embracing the need to maintain a particular set of beliefs about the world. This notion was inspired by hearing from people who feel that there has to be “something else” out there. This something else provides grander -- and therefore more meaningful -- explanations than just the mechanical versions provided by science. As an individual difference variable, nMyst may explain the striking variation in religious preference, susceptibility to religious conversion, and willingness to believe in the paranormal. The concept of nMyst may also make a contribution to ongoing research on the dynamics of meaning making (Heine, Proulx, & Vohs, 2006).

Clarification of the nMyst concept requires consideration of what counts as a mystery. Which mysterious beliefs have the motivational properties of a need? Two primary criteria for defining a mysterious belief are relevance to the person’s worldview and ambiguity in truth-value. A worldview is a set of assumptions and beliefs about reality that underlie people’s perceptions of the world around them (Koltko-Rivera, 2004). According to Rokeach (1973), worldviews include three types of beliefs; existential, evaluative, and prescriptive. Existential beliefs are beliefs about what exists in the world: They are capable of being true or false. Evaluative and prescriptive beliefs relate to values and whether an object or action is considered good or bad. Mysteries are existential beliefs assumed to have a truth-value. In the case of
nMyst, however, that value is ambiguous. Although believed to exist, there is no proof or clear way to gather proof of their existence. The most obvious example is belief in a god or gods: Deities either exist or they do not, but absolute proof seems impossible. Therefore such beliefs are assured to be stable. By setting such a high burden of disproof (following the adage, “absence of proof is not proof of absence”), such beliefs are well-protected.

Mysterious existential beliefs are an essential part of some people’s worldview. Beliefs are valuable insofar as they explain things about the world. The most valuable beliefs can explain a broad range of experiences and observations (e.g., the existence of God). Beliefs lose their value if they lose explanatory power (Preston & Epley, 2005), which accounts for why mechanistic explanations are often met with resistance. The claim that that humans exist because of evolution, not because God created them, undermines some of God’s explanatory power, thereby causing that belief to lose value. If someone who believes in ghosts or UFOs is told that all their observations could be explained by hypnosis or light reflections, this takes away from the explanatory value of believing in ghosts.

Valuable explanations tend to be functionalist in nature giving them intuitive appeal (Smith, Bruner & White, 1956). Children, when describing why things exist, give explanations based on purpose, e.g. “clouds are for raining” (Bloom, 2007). These explanations are described in philosophy as teleological which means they assume an intrinsic purpose to any natural phenomena, similar to purpose in human action. This intuitive bias can persist into adulthood. Evolution is an example of an explanation that does not specify purpose: Humans evolved through an ongoing series of environmental pressures and an unconscious drive to improve our fitness as a species. Because of its failure to provide a purpose or meaning to our existence, belief in evolution is not intuitively satisfying (Tracy, Hart, & Martens, 2011).
A related belief is in the separation of the physical body from a non-physical “soul”, not unlike differentiating between the brain and the mind (Bloom, 2007). Intuitive dualism is a basis for many religious and paranormal beliefs, e.g. the afterlife, ghosts, anthropomorphism of animals and objects, out-of-body experiences and so on. Reliance on intuitive thinking has been found to correlate with both religious and paranormal belief (Epstein et al., 1996; Wolfradt et al., 1999; Aarnio & Lindeman, 2007).

To many psychologists, it is not surprising that religious and paranormal beliefs overlap. Other commentators dispute that religious beliefs are analogous to paranormal beliefs (Hood, 1975). To pursue this issue, Aarnio and Lindeman (2007) used cluster analysis to separate types of believers: Skeptics, primarily paranormal believers, religious believers, and those who are both strongly religious and paranormal believers. The group that stood out was the Skeptics: They differed from all the other groups in scoring lower in intuitive thinking and higher in analytical thinking. Overall, the study provided further evidence to suggest that religious and paranormal beliefs are intuitive.

The specific beliefs described here are, of course, not the only possible mysterious beliefs. Any cherished belief that is held in a way which prevents disproof could be considered a mysterious belief. It is possible to hold even scientific theories as cherished mysterious beliefs and resist any evidence that contradicts them. Paranormal and religious beliefs are simply widespread and easily understood examples.

**Is Need for Mystery a Need?**

The traditional criteria for confirming a need included the demonstration of satiation, deprivation, etc.: Such behaviors are easily established in animal studies (Skinner, 1938). However, the term is now used more loosely to refer to a consistently strong motivational preference. Among these are needs for self-esteem, self-actualization, belonging, and a variety
of other concepts. Some modern theorists have proposed long lists of criteria for psychological needs based on making them analogous to physical needs (e.g., Baumeister & Leary, 1995). Many psychological needs have been proposed with little attempt to satisfy various traditional requirements for the label.

Sheldon et al. (2001) evaluated ten candidate needs using two criteria: Personally satisfying events should satisfy some psychological needs and the strength of the need should be related to the overall positive affect in response to the event. Among the ten were three from Deci and Ryan’s self-determination theory (1985): autonomy, competence, and relatedness. Based on their study, Sheldon and colleagues added the need for self-esteem. These two criteria seem to be satisfied by nMyst. It is not a stretch to imagine that, for those high in nMyst, hearing a new argument for one’s beliefs or joining a group supportive of mysterious beliefs would be satisfying and lead to more positive affect (see Gebauer & Maio, in press).

I believe that nMyst is also much better fitting with the concept of “need” than that of “want”. The literature on addiction has dealt extensively with the physiological distinctions between “wants” and “needs” (Love, James, & Willner, 1998). Wants are related to reward pathways, releasing dopamine and experiencing positive stimuli (Esch & Stefano, 2004; White, 1998). Needs are characterized more by an avoidance of negative stimuli and are related to serotonin (Barr et al., 2006; White, 1998). Being denied a want is a denial of pleasant experience and could be experienced as neutral; being deprived of a need is, by definition, a negative experience. In nMyst, the mysteriousness of the beliefs is needed to protect them from being disproven, this protection avoids the negative stimulus associated with worldview threat. The association between nMyst and worldview threat response will be addressed in the overall discussion in the context of the Meaning Maintenance Model (Heine, Proulx, & Vohs, 2006).
Also, the term “need” seems especially appropriate for Need for Mystery because of its motivational implications. Choice of this term emphasizes that people are especially invested in certain cherished beliefs: Certain human beliefs are held more strongly than others. This subset has been labeled core beliefs (Rokeach, 1973) or sacred beliefs (Haidt, 2007). One of my goals in isolating this variable is to better understand how people defend these cherished beliefs. The term “need” helps to emphasize the importance of these beliefs to those who hold them. Alternative labels such as value, goal, or preference, suggest more plasticity.

**Is nMyst a Personality Variable?**

A number of personality variables have been linked to needs (e.g., Murray, 1938). The prominent personality theorist, Edwards (1954), for example, construed all of his personality variables as needs. For those writers, the existence and strength of a personality variable implied a corresponding psychological need.

Confirmation of a personality trait tends to focus on a different set of criteria, namely, cross-situational consistency and temporal stability (e.g., Funder, 1997). The former typically involves confirmation that those who demonstrate the tendency in one domain show similar tendencies in other domains. The latter typically involves demonstration of a positive test-retest correlation.

Therefore, to establish nMyst as a personality variable, these two criteria will need to be established. For example, individuals with mysterious beliefs in one domain should show a similar tendency in other domains. Such individuals should also maintain these beliefs across time.

**Measuring nMyst: What should be included?**

To summarize, nMyst is a need to maintain beliefs that provide valuable explanations, are ambiguous in their truth value, and appeal to intuitive psychology. This need serves the function
of protecting the most cherished of human beliefs. I suspected, nonetheless, that people differ in the strength of this need.

To evaluate such individual differences, I sought to develop a questionnaire labeled the nMyst scale. But what content should be included? I began by circumscribing all the theoretical aspects relevant to the construct. A variety of facets emerged from an initial brainstorming with colleagues. We considered all possible individual differences between skeptics and those who believe in supernatural agents or mystical phenomena of any kind.

The six emerging topics -- Worldview rigidity, Failure of the Imagination, Intuition, First Person Experience, Animism, and Romanticism -- guided the initial question writing:

1) *Worldview rigidity* was measured with items expressing a dislike for information that undermines one’s personal worldview. Examples include avoiding secrets of how magic tricks are done, or insisting that there are things that people should not try to understand completely.

2) *Failure of the imagination* relates to the idea of naïve realism (Ross & Ward, 1996). Naïve Realism is the idea that any personal experience is assumed to represent reality until one acquires contradictory evidence. Even if the personal experience seems unlikely, such as having an out-body-experience or “seeing” a ghost. The notion is exemplified by the question “I find it hard to believe in things I can’t picture”.

3) *Intuition* is a preference for emotionally-appealing (or common sense) over scientific explanations. Because it involves a source outside of objective information, intuition serves to define the common mysteries. The notion is exemplified in the item “I trust my intuition over scientific evidence”.

4) *First person experience* is an extension of the notion of naïve realism: Not only do I trust my eyes, but I trust others’ eyes as well. If someone has had a personal
experience, either spiritual or paranormal, their account is persuasive enough to be held as evidence that experience was real.

5) Animism is the tendency to assume deliberate agency in non-humans. It is part of a more general tendency towards anthropomorphism of animals and inanimate objects. The tendency has been theorized to be a precursor to all human religions, and possesses strong intuitive appeal (Norenzayan, 2010).

6) Romanticism is the tendency to glorify everyday experiences as spiritually meaningful. Examples include feelings of awe when gazing at the stars or the joy in being transported by music. By contrast, scientific explanations seem dry, boring, and sometimes undignified.

Worldview rigidity and failure of the imagination were based on arguments often heard when religious beliefs are challenged. They often involve the rejection of concrete explanations in favor of teleological explanations. Intuition is clearly relevant because it energizes the appeal of teleological explanations: They just feel correct. A reliance on intuition rather than critical thinking has been found to relate to paranormal belief in earlier research (Alcock & Otis, 1980; Gray & Mill, 1990; Wierzbicki, 1985).

First person experience was included because it can be a part of intuitive reasoning. It feels intuitive to rely on your own experiences as data points. Anecdotal evidence is generally accepted as a proof of an argument outside of scientific circles. When learning the scientific method students must be explicitly taught to ignore such evidence, but without training most people will accept it.

Animism is part of nMyst in that it is, itself, a commonly used teleological explanation. Assigning agency to inanimate objects and forces allows one to explain phenomena as an intentional action. Perceiving agency and reasoning about intention are fairly automatic
processes that are an important part of social functioning. Interpreting one’s environment using these heuristics should be intuitive, and cultural evolution theories suggest that this is one of the pre-cursors to religious belief (Atran & Norenzayan, 2004). Romanticism is the more affective preference component of nMyst. It encompasses a tendency to experience awe and enjoy inexplicable or spiritual events. It is similar to concepts such as Mysticism (Hood, 1975) and Absorption (Glisky, et al., 1991), which will be examined in greater detail in Study 5.

These six facets were assumed to be indicators of the latent variable postulated by the theory of nMyst. However, preliminary statistical analyses revealed that they clustered into two broad factors. One factor is primarily defined by the Worldview Rigidity and Failure of the Imagination facets; the other is defined primarily First Person Experience, Animism, and Romanticism. Intuition items tend to relate to both factors. The study in which these factors were determined will be described later. I will give a detailed description of the two factors here to better orient the reader to the theoretical arguments.

These two broader factors made eminent sense in the context of other belief research incorporating both religious and paranormal belief. They provide a more manageable and coherent organizational framework for the concept of nMyst. Based on the content of the items, the axes were labeled Anti-Science (AS) and Pro-Magic (PM). The factors can be interpreted as different strategies¹ for maintaining mysterious beliefs. Mysterious beliefs provide comfort in their stability. These strategies satisfy a comfort-seeking motivation to protect cherished beliefs (Rokeach, 1973).

¹ My use of the word ‘strategy’ is used for simplicity and does not imply conscious control.
Anti-Science

Anti-Science (AS) is the more strongly defensive strategy for maintaining mystery. The facet of Worldview Rigidity is very important to this strategy. Those high on AS protect their worldview by directly rejecting potentially disconfirming or devaluing evidence. They express a distrust of scientific evidence and empiricism, in general, preferring anecdotal or intuitive evidence that supports their existing existential beliefs.

The value of existential beliefs lies in their explanatory ability (Preston & Epley, 2005). People who are high in AS want to keep their explanations mysterious and thus more valuable. Again using the example of evolution, it is generally understood by scientists that it is technically possible to find evidence that does not fit with evolution as currently conceptualized and the theory will have to change to accommodate this new evidence. This is the peril of accepting scientific explanations; they are susceptible to change or even disproof. Thus those high in AS choose to maintain explanatory beliefs that are not susceptible to evidence, their value is protected because the only thing needed to make something true is their personal belief that it is true, which is far more stable.

As one would expect, Anti-Science is highly related to religiosity. Religious belief bases all mystery on the existence of a particular deity or deities and so, to maintain mystery that belief must be protected. The Anti-Science strategy is particularly important when the mysterious worldview hinges on a single belief. The AS strategy is conceptually related to types of motivated cognition, or ways of thinking that can protect a person from anxiety. Examples of motivated cognition are conservative attitudes included as part of the System Justification Theory (Jost & Banaji, 1994). These attitudes include Authoritarianism (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950), or Belief in a Just World (Lerner, 1980). System Justification Theory has been shown to be a form of worldview defense and is dependent on trust
in the system being part of a person’s worldview to activate a system justifying response (Banfield, et al., 2010). The need for worldview structure is also a type of motivated cognition and operationalized as either Personal Need for Structure or Need for Cognitive Closure (Neuberg & Newsome, 1993; Webster & Kruglanski, 1994). These constructs describe a need to have any firm belief on a particular topic and have been found to relate to political conservatism and religious belief (Duriez, 2003; Matthews, Levin, & Sidanius, 2009).

**Pro-Magic**

The Pro-Magic strategy is expansive rather than defensive in regards to particular beliefs. People who employ the Pro-Magic strategy are still defending the general mysterious worldview, but do it by adopting many varied existential beliefs rather than a few that must be jealously guarded. Those high on Pro-Magic tend to see every day experiences as mysterious, like listening to music or looking at the stars. They are also more likely to hold paranormal beliefs, such as belief in ghosts or extrasensory perception.

The existential beliefs maintained using the Pro-Magic strategy are varied enough that if a particular one becomes devalued, it can be replaced by a different mysterious belief. They are not necessarily against science either (Irwin, 1993). In fact, many people who believe in the paranormal or extraterrestrials actively seek scientific confirmation of their beliefs in groups like the Princeton Engineering Anomalies Research Group (PEAR) or the Search for Extraterrestrial Intelligence program (SETI). People such as ghost hunters also look for specific evidence for their beliefs. Rather than denounce scientific findings as unintuitive, they recognize the boost in credibility to their beliefs gained by applying scientific methods. One important thing to note, though, is that they rarely let negative results in these investigations temper their beliefs. They tend to choose to believe only positive findings and show selective learning in the face of disconfirming evidence (Russell & Jones, 1980).
Pro-Magic is conceptually related to Awe-proneness, or the tendency to experience awe. Awe is an emotion that is theorized to occur in the presence of something experienced as “vast” which requires accommodation (Keltner & Haidt, 2003). In this context “vast” refers to an expansion of the usual frame of reference, from more literal instances like looking up into space to more metaphorical things like a work of art “expanding perception”. The expansion then has to be accommodated in the self-concept or schema.

The tendency to feel awe is then associated with an ability or willingness to change or expand schemas (Shiota, Keltner, & Mossman, 2007). In order to use the Pro-Magic strategy, a person’s schema of “mystery” must be able to expand and be flexible. An experience of awe can expand the scope of mysterious beliefs. Those who are high in Pro-Magic tend to experience more awe in response to common experiences and thereby continuously add evidence to their mysterious worldview.

Unlike Anti-Science, Pro-Magic is unrelated to religiosity. The mysteries contained in religious teachings are fairly specific and all interrelated. The belief in God is the primary belief on which all mystery is built. This is an unstable worldview system which needs constant vigilant protection, i.e. Anti-Science. The Pro-Magic strategy can include religious belief, but it expands on that by incorporating other mysteries leading to a more stable system. Because Pro-Magic can incorporate mysteries from anywhere, people who employ that strategy are no more or less likely to be devoutly religious.

The two strategies could be thought of as potentially complementary. Anti-Science is reactive and protects beliefs from threats. Pro-magic is pre-emptive; providing a broader range of beliefs and softening the sense of unease caused by a threat to one belief. The particular strategy a person employs may be learned, while the general need for these beliefs is perhaps more intrinsic.
The Quadrants

Because the AS and PM factors are both fundamental to the overall construct of nMyst, I hypothesize a positive correlation. Nonetheless, the strategies are conceptually distinct and should not be empirically redundant. For example, AS should correlate with religiosity while PM should not. PM should be positively correlated with belief in the paranormal whereas AS should not. Although it is possible to be high in AS and low in PM or vice versa, it is also possible to be high or low in both using the two strategies in tandem or not needing mysterious beliefs at all. Figure 1 provides an illustration of the quadrants and how each would theoretically be implemented.

If one is high in nMyst then using both strategies would provide the most efficient protection. Protecting deeply held beliefs while bolstering them with more numerous beliefs that are less central would provide a very stable mysterious worldview, the way a four-legged table is more stable than a pedestal. Those scoring highest on nMyst would appear in the high-high group. Those who score lowest in nMyst have no need to employ any mystery protecting strategies and so are naturally low on both.

Although I will not be elaborating on the quadrants in the analyses, they serve as a graphical reminder that the strategies are not mutually exclusive. A comparison of different religious doctrines should give some insight into how the two strategies are used together in some belief systems, whereas other religions may encourage the use of only one.

Overview of the Four Studies

The practical goal of this dissertation is to develop the nMyst measure and provide evidence for its validity as measure of the theoretical construct. To this end, I have conducted four studies aimed at establishing the reliability (both internal consistency and test-retest), and establishing gender and ethnic norms. I will compare nMyst to measures of related constructs to
establish convergent and divergent validity. Each study includes a variety of hypotheses based on the theoretical notions outlined in this introduction.

Study 1

Development of the nMyst measure began with an initial item pool of 32 items addressing the six facets discussed in the introduction. Study 1 was the initial evaluation of these items to determine the actual factor structure and to eliminate poorly performing items. In addition, several established measures and demographic questions were included to begin to build the nomological network in which nMyst is embedded. These included the Big Five personality traits (McCrae & Costa, 1987), Personal Need for Structure (Neuberg & Newsom, 1993), and tendency to experience awe (Shiota, Keltner, & Mossman, 2007).

I also included a set of items asking participants to rate the plausibility of a variety of paranormal phenomena. Recall that the theory behind nMyst suggests a trait-like tendency to believe in mysterious things. The most obvious way to test this is to see if nMyst scores actually predict belief in various concepts that we would consider mysterious.

Study 2

In Study 2, I obtained a very large community sample allowing for a thorough statistical analysis of the measure as well as population norms. I used these data to perform a series of confirmatory factor analyses (CFA) to follow up on the initial exploratory factor analysis (EFA). I also ran a factor analysis comparing European heritage and Chinese heritage participants. Item Response Theory analysis was used to determine the functioning of individual items and look for differential function of the items between ethnic groups.

The sample was very diverse allowing for education and cross cultural comparisons. Many different religious beliefs were represented allowing me to investigate whether different
religions use the two nMyst strategies differently. Correlations with the Big Five are also included.

This dataset also included many participants who had experiences with religious conversion. Overall, I hypothesized that nMyst scores will more resemble the average of the post-conversion religion than the pre-conversion religion. I expect this to be the case not only for those who reject religion, but also people who have a more traditional conversion from one religious faith to another.

**Study 3**

Study 3 evaluated the relationship between nMyst and several political attitudes: The Moral Foundations Questionnaire (Graham, Haidt, & Nosek, 2009), Belief in a Just World (Lerner, 1980), and Right Wing Authoritarianism (Adorno et al., 1950). Conservative political attitudes were expected to correlate positively with AS, as they are usually also related to religiosity and motivated cognition. PM was expected to correlate negatively with Belief in a Just World, and possibly with some of the moral foundations.

I included a measure of free will belief (The FAD-Plus; Carey & Paulhus, in press; Paulhus & Carey, 2011) because that concept fits the definition of mysterious beliefs as being functional rather than mechanistic. Believing in free will requires the belief that there is something special about human agency and that it derives, not from the brain or environment, but some undetermined control center. This common assumption implicates the existence of a soul, or at least a mind, operating independently of the body. It is likely that both PM and AS will correlate with belief in free will because it is related to more traditional religious beliefs as well as a more general orientation towards teleological explanations.
Study 4

In Study 4, I addressed some issues of discriminant validity and measured the test-retest reliability of the nMyst measure. I correlated nMyst with a number of constructs which have been suggested to have conceptual similarity. Schizotypal personality disorder (Raine, 1991), Absorption (Tellegen & Atkinson, 1974), and Mysticism (Hood, 1975), are all related to the PM subscale as they capture different aspects of spiritual experience or magical thinking. I expected that these constructs should be somewhat related to PM, but should not have a high correlation as PM is designed to capture a broader construct.

Summary

This introduction has served three broad purposes. One was to introduce and elaborate on a new theoretical construct that I have dubbed nMyst. nMyst as a construct describes a need for certain types of beliefs and explanations. These beliefs must be unfalsifiable and provide functionalist explanations for observed phenomena. Primarily, these beliefs are either paranormal or religious in nature.

The second broad purpose of this introduction was to detail my plan for developing and validating a measure of nMyst. The theory suggests a number of criteria that should be met by the measure if it is, in fact, a valid measure of nMyst. I provided an outline of four empirical studies designed to evaluate the structure, reliability, and validity of the nMyst scale.

To anticipate, these four studies reveal a coherent nomological network for nMyst. Scores on the nMyst scale are shown to predict trait-like tendency to hold a broad range of mysterious beliefs. Associations with established individual difference measures provided appropriate support for both convergent and discriminant validity. Other data support the predicted role of nMyst beliefs in religious conversion: A mismatch between beliefs and nMyst leads people to change in a more consonant direction. The reliability of the nMyst measure was
supported using both internal consistency and test-retest methods. These findings were extended cross culturally and expanded using IRT methods.

Together, the analyses presented in this thesis indicate that the nMyst scale is valid measure of the intended construct and is stable across time and multiple samples. Availability of the new instrument should facilitate future research on the role of nMyst in protecting cherished beliefs.
Chapter 2: Item Development

To begin to investigate Need for Mystery, I designed a scale to measure individual differences in the construct. I started by creating an initial item pool of 32 items addressing the six facets as discussed in the introduction, these item can be found in Appendix A. Study 1 tested these items to determine the actual factor structure and to eliminate poorly performing items. No specific subscale hypotheses were made a priori because the general structure of the scale was not yet known. I hoped to find a latent factor which captured the common element of the six facets. In this sense, a single factor would have been ideal though unlikely given the breadth of topics included in the scale. I chose to use Parallel Analysis to guide my choice on the number of factors present in the data. Parallel Analysis (PA) is generally more reliable than methods such as eigenvalue-greater-than-1 or scree plot methods (Wood et al., 1996; Zwick & Velicer, 1982, 1986). Choosing all factors with eigenvalues greater than one can often lead to including trivial factors (Zwick & Velicer, 1986). Scree plots can be difficult to interpret as the “elbow” is often unclear and subjective and even experts are not very reliable in their interpretations (Crawford & Koopman, 1979; Streiner, 1998). PA is more directly empirical and is useful in cases in which there is no strong theoretical hypothesis to guide the interpretation of the scree plot or eigenvalues. I will describe the technical details of PA and the MAP in the results section.

In addition, some basic scales and questions were included to begin to establish the validity of the nMyst as a new construct. Establishing the nomological net for nMyst requires beginning with basic personality traits, like the Big Five, as well as situating it within the framework of other possibly related individual differences. Personal Need for Structure measures a tendency to reduce information inputs from the world through the use of preexisting structures. PNS manifests as a reliance on schemas and stereotypes and a tendency to assimilate
new information rather than accommodate new information. PNS also includes items addressing Need for Cognitive Closure which describes a preference for clarity and avoidance of ambiguity. It may seem from this description that nMyst might be negatively correlated with PNS; however, I hypothesized that they would be positively related. The use of teleological and unfalsifiable explanations leads to greater structure and certainty in one’s worldview. As I outlined in the introduction, these types of beliefs are actually more stable than concrete explanations that may be difficult to find in the first place and subject to revision. PNS is positively related to other constructs such as Authoritarianism, Dogmatism, Intolerance of Ambiguity, and Rigidity, which are all conceptually related to the Anti-Science factor of nMyst.

The spiritual aspect of nMyst also has its own distinct nomological net separate from the motivated cognition aspect. In the first study, I have included a measure of tendency to feel awe as a criterion variable. Awe in psychological literature has been operationalized as a response to stimuli that are “vast” in some way, not necessarily size, and require accommodation (Keltner and Haidt, 2003). Awe provoking stimuli must be beyond the perceiver’s frame of reference and require them to expand their understanding of the world. Example of awe provoking stimuli can be related to nature, like the ocean or a mountain, art or music, or religious experience. Awe is associated with feelings of decreased attention to the self and a greater sense of one’s connection to the universe (Shiota, Keltner, & Mossman, 2007). It was also found to be negatively correlated with Need for Cognitive Closure which is addressed as part of Personal Need for Structure. Awe is often related to religious experiences and some of the Romanticism items in the nMyst scale describe scenarios that could be interpreted as awe, e.g., “I sometimes get a shiver thinking about mysterious things.”

Another important part of establishing scale validity is showing that it predicts observed variables as well as other latent constructs. In the case of nMyst, it should obviously predict the
theorized teleological beliefs. These include both religious and paranormal beliefs as well as conspiracy theories and other non-spiritual “mysteries”. I included a list of several of these beliefs and asked participants to rate them from “Definitely Not Real” to “Definitely Real”. Specific items came from six categories; parapsychology (e.g., telekinesis), paranormal (e.g., ghosts), conspiracy theories, extraterrestrials, and religious based “magic” or miracles (e.g., the effectiveness of prayer).

**Method**

The initial set of items contained 32 items designed around the six proposed facets. These items were administered online to a student sample. Ethnicity was not collected due to a computer error; number of years in North America was used as a proxy variable using only participants who had been in North America for at least 12 years. I was left a sample of \( N = 217 \) (71% Female). Also included were the Big Five Inventory (BFI; John & Srivastana, 1999), Personal Need for Structure (PNS; Neuberg & Newsome, 1993), and Awe (Shiota, Keltner, & John, 2006). Reliabilities for the BFI subscales were all high: Extraversion, \( \alpha = .83 \); Agreeableness, \( \alpha = .81 \); Conscientiousness, \( \alpha = .81 \); Neuroticism, \( \alpha = .84 \); Openness, \( \alpha = .78 \). PNS and Awe also had good internal consistency reliability, \( \alpha = .80 \), and \( \alpha = .75 \), respectively.

Participants were also asked to rate 17 different religious and paranormal beliefs on a 7-point Likert scale from “Definitely not real” to “Definitely real”. A list of the items can be found in Appendix B. Demographics such as gender and religiosity were also collected.

**Results**

**Factor structure.** To determine the number of factors, I used a parallel analysis (PA) and the MAP test. PA is suggested to account for non-normality or violations of continuous variable assumptions caused by using a Likert scale. The Parallel analysis procedure compares the eigenvalues extracted from the data to eigenvalues in a randomly resampled set of the data;
this is a similar procedure to bootstrapping in regression. The output for a PA lists the eigenvalues for the data compared to the randomly sampled eigenvalues and the number of eigenvalues in the data which are greater than the random eigenvalues indicate the number of non-trivial factors. The parallel analysis indicated that the first three eigenvalues were well above chance, but the fourth was only slightly higher, 1.83 compared to 1.62. Since even the PA was not entirely conclusive, I followed up Velicer’s MAP test (1976). The MAP test uses Principle Components Analysis to determine how many factors must be removed from the correlation matrix of the data to minimize the remaining off-diagonal partial correlations. This results in the number of factors which are still accounting for systematic variance rather than noise. The MAP test confirmed that only 3 factors were non-trivial.

A three factor EFA was run using ML extraction and Oblimin rotation. The loadings for this analysis are shown in Table 1. It is clear from the table that a number of the items had low loadings on all three factors. The highest loading items from Factor 1 were from the Worldview Rigidity, Failure of the Imagination, and Intuition facets. On Factor 2, the high loading items included Animism and First Person Experience. Factor 3 had only 3 items with loadings greater than .4 and all were from the Romanticism facet and crossloaded on other factors.

Items with loadings below .3 were dropped and another three factor EFA was performed on the new item set. Three items had loadings below .3 and they were dropped from the subscale calculations. The loadings for the final item set are shown in Table 2. I then tested the alpha reliability of the three subscales. Factor 1 had an alpha reliability of .83 for 8 items. The reliability for Factor 2 was much lower, $\alpha = .72$ for 6 items, and of course the reliability for Factor 3 was quite low with only 3 items, $\alpha = .62$. Because Factor 3 only accounted for three items it did not seem valid to treat it as a separate factor in subsequent analyses. Based on correlations between the three factors, I chose to combine it with Factor 2. The subscale
intercorrelation was significant ($r = .24, p < .001$) and the item content was conceptually similar. I did not drop the factor because I felt that Romanticism was an important part of the construct and dropping the items would damage the scale validity. The new Factor 2 had an alpha reliability of .70 with 9 items. The correlation between Factor 1 and the new Factor 2 was .27. These factors were given the labels Anti-Science and Pro-Magic.

**Validity.** No significant gender differences were found on either AS or PM. The BFI results showed that AS was positively correlated with agreeableness, $r = .18, p < .01^2$. PM was positively correlated with neuroticism and openness to experience, $r = .16, p < .05$, and $r = .17, p < .01$ respectively. PM was correlated positively with a tendency to experience awe, $r = .25, p < .001$. PM was negatively correlated with PNS, $r = -.14, p < .05$.

Anti-Science (AS) was found to correlate highly with religiosity, $r = .64, p < .001$. Pro-Magic (PM), however, was unrelated to religiosity, $r = .07$. The paranormal belief items were aggregated into a single composite which had an alpha reliability of .86. PM was correlated with the total composite, $r = .35, p < .001$. However, some items from the PM subscale had some overlapping content with the paranormal items: items 18, 20, and 29 specifically mention ghosts, telekinesis, and clairvoyance, respectively. After removing these three items from the PM subscale, the overall correlation dropped to $r = .18, p < .01$, but it was still significant. AS was not correlated with the entire composite, however it was positively correlated with the parapsychology items, $r = .18, p < .01$, and negatively correlated with belief in aliens, $r = -.15, p < .05$.

^2 All significance tests reported in this thesis are two tailed.
Discussion

The EFA results suggested a three factor solution in which the romanticism facet loaded on its own factor separate from Animism and First Person experience. However, these two factors on their own had low alpha reliabilities and a significant intercorrelation. A factor combining these two had much better reliability with all items improving the alpha estimate. The scale is designed to measure a fairly broad construct and it is unsurprising that items measuring only one facet would have higher correlations with each other than the larger construct. However, I am more interested in the theoretical latent variable that is defined by the overlap among the facets than in their individual behavior. Therefore, I feel that combining these factors is more valid theoretically than treating Romanticism as a separate factor. For the rest of the analyses I used the Anti-Science and Pro-Magic subscales separately.

The results of the big five personality traits were interesting. The correlation between AS and Agreeableness suggests that perhaps these people see scientific explanations as somehow causing conflict by questioning currently held beliefs. AS was not negatively correlated with Openness to Experience; however, PM was positively correlated with Openness which is also consistent with idea of being open to new beliefs and suggests a bit of non-conformity to the PM factor. PM being correlated with Neuroticism was unexpected, but perhaps having a very strong belief in the supernatural could lead to anxiety. It is also possible that people high in PM have a tendency to be overly dramatic and this leads to stronger emotional reactions and greater variability.

PNS also did not fit my hypothesis. It was negatively correlated with PM rather than positively correlated with AS, as predicted. It does make sense for PM to be negatively correlated with PNS. PM is a more unstructured type of belief that is unrelated to the rules and framework provided by specific religions. It is possible that the lack of relationship between
PNS and AS is a results of range restriction on PNS in the student sample, though unfortunately I did not follow up by including PNS in any of the community sample studies. It may also be the case that the stability of beliefs provided by the AS strategy are not related to a general need for structure, but simply a consequence of the types of beliefs associated with that strategy. Awe was positively correlated only with PM which fits with the hypothesis that Romanticism is somewhat based on interpreting experiences of awe as spiritually meaningful. AS was not correlated with Awe, however religious awe specifically is not included in this scale.

The paranormal belief items were all moderately correlated with PM which supports the basic theory that we are tapping a latent factor that underlies all these types of beliefs rather than measuring paranormal beliefs themselves. The correlation held even when items with explicit paranormal content were removed from the PM subscale.

Belief in extraterrestrials is an interesting case because it can be argued from a purely scientific standpoint that they exist; however it has also been adopted into the canon of paranormal beliefs. Thus those types of beliefs were positively correlated with PM and negatively correlated with AS.

The one item measure of religiosity was correlated positively with AS and unrelated to PM. The general theory of AS is that it is based on a more rigid and structured type of spirituality. In contrast, PM is more general. However, PM is unrelated to religiosity rather than negatively correlated, so incorporating PM into a religious worldview is certainly possible. This finding led to my hypothesis that different religious beliefs may incorporate PM, more or less, depending on their belief structure and rituals.
Chapter 3: Psychometrics, Demographics, and Religious Conversion

So far, my analyses of the nMyst scale have been restricted to student data. Study 2 sought to extend its generalizability to a larger community sample. To this end, I collected a large dataset on Facebook.com: It included standard personality variables as well as demographics and current and childhood religious affiliation.

Of special interest was confirmation of the scale structure emerging from Study 1. Accordingly, I conducted a series of confirmatory factor analyses (CFAs) following up the exploratory factor analysis from Study 1. In a CFA the proposed factor structure is compared to that found in the actual data and the misfit of the data to the proposed factor structure is quantified in several fit test statistics. I used CFA to empirically test the appropriateness of the model suggested by the EFA. Once I established a fitting general model, I ran a multiple groups CFA (MGCFA). MGCFA tests the data from two different sample groups simultaneously to determine whether the fit of the proposed structure is comparable across the two groups. I chose to compare European heritage participants and participants who self-identified as Chinese. Chinese heritage participants were the second largest ethnic group in sample allowing for equal Ns in the analysis. It was also important to establish the scale structure in an Asian heritage group in order to justify combining European and Asian heritage participants from the UBC student samples.

In addition to the CFA, I applied Item Response Theory (IRT) methods to assess the behavior of the scale on the item level. IRT analyzes each item in terms of the probability of a particular answer for a person at a certain latent trait level (Embretson & Reise, 2000). In the case of Likert items, it is based on the probability of each response given the person’s latent trait level. For each item, the trait level for which each response category (each number from 1: Strongly Disagree to 5: Strongly Agree to) is selected at a probability of .5 is called the
“difficulty” parameter for that response category on that item. Additionally, the extent to which each item can distinguish between different trait levels is estimated as the “discrimination” parameter. Discrimination is related to latent variable loadings in factor analysis, in that items with strong loadings have higher discrimination.

Though IRT was developed for use with categorical items, it is still useful in cases with Likert items which are actually categorical and may not behave as completely continuous variables, especially at only 5 categories (Rhemtulla, Brosseau-Liard, & Savalei, 2012). This analysis can give a more detailed account of the range of the latent trait being measured as well as how the response categories are being used. It also shows the differences in measurement error at various trait levels, these differences are reflected in the reliability of correlations. Finally, IRT provides an estimate of the standard error of estimate for the scale at each point on the latent trait: A rescaling of the standard error is reported as Test Information.

Item response theory also allows for a test of Differential Item Functioning (DIF). DIF compares the behavior of an item in two groups to identify items that are behaving differently for different groups. For example, some items may perform more poorly for non-native English speakers: They may be more difficult or have a different interpretation. In this study, I compared European and Chinese heritage participants to determine which items have significant DIF for these two groups.

I also sought to compare different ethnic groups on their levels of nMyst. My conception of nMyst is grounded in a western view of religion as separate from culture and subject to individual choice. If that notion is indeed uniquely western, it may not have validity in other cultures. In addition to structural analyses on the Chinese heritage subsample, I examined scale reliability and mean differences in the 12 different ethnic groups represented.
Religious Conversion

Understanding religious conversion was one of my primary interests when constructing nMyst. To investigate this phenomenon, I collected both childhood religious affiliation and current religious affiliation. There were a significant number of participants who had true conversion experiences in the traditional sense of moving from one religion to another, not just leaving religion entirely.

I hypothesize that religious beliefs will change in response to a mismatch between a person’s nMyst levels and their current beliefs. Of course, it is difficult to determine temporal precedence of nMyst attitudes and religion. Indoctrination into a religion that specifically denies scientific findings may cause a reliance on the AS strategy, making the religious belief causal. However, does that mean that the need to maintain mystery is also caused by childhood religious teachings? It could be that nMyst is a stable personality trait and the protection strategy is determined by the particular belief.

Evidence for the temporal precedence of nMyst could be found by studying religious conversion. Religious conversion has been conceptualized as a multi-stage process which involves a crisis of faith, exposure to a new faith, and eventually devotion to the new belief (Rambo, 1993). Conversion is a voluntary and meaningful life change. Some have reported that the process of religious conversion caused a profound change in their personality and behavior, being “born again” as it is sometimes called.

If it is true that voluntary religious change can affect personality then it is plausible that religious belief takes precedence in the case of nMyst. However, a review of studies on religious conversion suggested that personality is only changed at the levels of Characteristic Adaptations and Life Narrative, rather than traits (Palutzian, Richardson, & Rambo, 1999). Characteristic Adaptations are a person’s goals, concerns, attitudes, and personal strivings (Costa & McCrae,
Life Narrative is how people organize and describe their identity and define their purpose in life (McAdams, 1994). Because it is a trait, I hypothesize that the baseline need for mystery would eventually become fixed and would not change as results of conversion (Costa & McCrae, 1994).

These data are cross-sectional, so I cannot directly test my hypothesis. However, I can test the proposition that people change religious affiliation to one that is better aligned with their nMyst scores by comparing converts to non-converts on both their childhood and chosen religion. I expect that converts will, on average, have scores more similar to the averages of non-converts in their chosen religion than to non-converts in their childhood religion.

**Method**

Participants were recruited on Facebook as part of an application called “My Type” ($N = 15,630$, 60% Female, 63% European heritage, 11% Chinese heritage, $M_{age} = 25.60$). A 23-item version of NMyst was included along with a large number of demographics, including childhood and current religious affiliation. A subset of participants also completed a Big Five personality measure based on the IPIP-50 ($N = 9,857$).

**Results**

**Confirmatory Factor Analysis.** I conducted a series of confirmatory factor analyses using the program EQS (Bentler, 2004) using robust Maximum Likelihood estimation. The fit indices for all the reported models can be found in Table 3. The initial model was a three factor version with 21 items and correlated factors with no cross loadings or correlated errors (Model 1). Factors were identified by fixing one loading at 1.0. The model was run on a subsample of 500 participants. Because the normalized kurtosis estimate was 4.50, robust fit statistics are reported for all models using these data including the Satorra-Bentler scaled chi-squared (Satorra & Bentler, 2001).
The initial Model 1 showed a poor fit based on several indices, Satorra-Bentler $\chi^2(186) = 613.58, p < .001, CFI = .85, TLI = .83, RMSEA = .07 [0.06 - 0.08]$. However, the loadings were generally high and the average off-diagonal standardized residual was only .06. Lagrange Multiplier Tests (LMTEST) were used to determine if any additional parameters should be added to improve model fit. For Model 2, five suggested cross-loadings were added, four between Factors 1 and 2 (Items 12, 16, 17, and 22), and one between Factors 2 and 3 (Item 27). The Factor 1 and 2 cross-loadings were generally small (< .30), and for Items 16 and 17 they were actually negative. The fit was improved significantly in Model 2, according to a chi-squared difference test, but still did not meet standard fit criteria, Satorra-Bentler $\chi^2(181) = 482.27, p < .001, CFI = .90, TLI = .88, RMSEA = .06 [0.05 - 0.07], \chi^2_{\text{diff}}(5) = 131.31, p < .001$. The average off-diagonal residual was .05. The correlation between Factors 2 and 3 in this model was very high, $r = .69$, suggesting that a two factor model might be more appropriate.

I tested a two factor model (Model 3) in which all Factor 3 items were changed to load on Factor 2. I initially tested a model without any cross-loadings to be comparable to the unadjusted 3 Factor model. The fit was poor by all standards, Satorra-Bentler $\chi^2(188) = 653.25, p < .001, CFI = .84, TLI = .82, RMSEA = .07 [0.07 - 0.08]$. However, it was comparable on most indices to the unadjusted 3 factor structure in Model 1, though the chi-squared difference test was significant; $\chi^2_{\text{diff}}(2) = 39.67, p < .001$. As a follow up, I tested a 2 Factor model with the cross-loadings found between Factors 1 and 2 in Model 2 (Model 4). The covariance between Factors 1 and 2 was still estimated. Model 4 did not fit the data by the chi-squared test and the chi-squared was significantly worse than Model 2, Satorra-Bentler $\chi^2(184) = 529.61, p < .001, CFI = .88, TLI = .86, RMSEA = .06 [0.05 - 0.07], \chi^2_{\text{diff}}(3) = 47.34, p < .001$. The average residual was still less than .05. Most of the loadings were still quite high, with one item on Factor 2 dropping below .3, which also had a fairly low loading in the 3 factor model (Item 22).
I decided that Model 2, the 3 factor model with cross-loadings, was likely the superior base model due to the fit statistics and concordance with the previous EFA. To improve the fit of the model, I included possible correlated errors in the LMTEST to see if the fit could be improved through adding some within factor error correlations (Model 5). The test suggested releasing several error terms in Factor 1. The four most significant correlated errors were for items 8 and 9, 13 and 21, 13 and 8, and 22 and 23. Though the model still did not fit by the chi-squared, it did fit by the combined criteria of $CFI > .95$ and $RMSEA < .05$ (Hu & Bentler, 1999), Satorra-Bentler $\chi^2(177) = 361.89, p < .001$, $CFI = .95$, $TLI = .93$, $RMSEA = .05 [.04 - .05]$. This was also a significant improvement in fit over Model 2 by the chi-squared difference test, $\chi^2_{diff}(3) = 120.38, p < .001$.

To follow up on the model adjustments, I ran the 3 factor model with cross-loadings on a second subsample of 500. The model was identical to Model 5 with correlated errors. The normalized kurtosis estimate was quite significant in this sample, $Z = 10.08$, so robust fit estimates are reported. The fit was comparable to the fit in the previous dataset, though it still did not met the chi-squared criteria$^3$, $\chi^2(177) = 415.91, p < .001$, $CFI = .92$, $TLI = .90$, $RMSEA = .05 [.04 - .06]$. The standardized off-diagonal residual was .04. Loadings were comparable to Model 5 and better in some cases. This model with loadings is shown in Figure 2. The factor correlation between Factors 2 and 3 was $r = .69$, which provides further support for the combining of two factors into the nMyst subscale when forming the composites for ease of theoretical interpretation.

$^3$ The repeated poor fit for the $\chi^2$ test is not a serious concern. It depends on sample size and has been found to be far too stringent for most personality constructs. Other criteria (such as those outlined in Hu & Bentler, 1999 or Yuan, 2005) are more appropriate for this context (Hopwood & Donnelan, 2011).
**Multigroup CFA.** In addition to the normal CFA analysis I used a multigroup CFA to compare European and Chinese-heritage participants. A dataset composed of 654 European heritage participants (self-identified as “White”) and 575 Chinese heritage participants (self-identified as “Chinese”) were compared on Model 2 to test the similarity of the factor structure between groups. I did not include a mean structure comparison because I have already shown that the means are significantly different between the two groups. The multigroup model did not fit suggesting that the loadings structure is significantly different between the two groups, Satorra-Bentler $\chi^2(362) = 1119.47, p < .001, CFI = .86, TLI = .84, RMSEA = .06 [.05 - .06].$

Though the poor fit suggests that the factor structures are not identical, the actual factor loadings were still comparable in the Chinese heritage sample. The average loading for Factor 1 was .60 in the European sample and .50 in the Chinese heritage sample. The average loadings for Factor 2 were almost identical, .52 compared to .50, and Factor 3 actually had higher loadings in the Chinese heritage sample, .43 compared to .49. The factor correlations were much higher in the Chinese heritage sample, suggesting that perhaps nMyst is a more unitary construct in other cultures.

To follow up on this analysis I attempted to find a fitting model for the Chinese heritage dataset. I started with a two factor rather than a three factor model, given the extremely high correlation between Factors 2 and 3 (.85) in the MGCFA. I also removed the cross-loadings from the European dataset model and ran LMTESTs for releasing both cross-loadings and errors. The normalized kurtosis estimate was $Z = 18.10$, so robust fit statistics are reported. The model was not a good fit to the data, Satorra-Bentler $\chi^2(169) = 472.00, p < .001, CFI = .83, TLI = .81, RMSEA = .06 [.05 - .06].$

A second model allowing for both cross-loadings and additional correlated errors based on the LMTEST. Items 12 and 22 were allowed to cross-load on Factor 2 which is consistent
with the European structure. Error covariances were estimated between Items 13 and 8, 12 and 9, 22 and 23, 20 and 27, and 31 and 32. Also, Item 25 was dropped due its very low loading (.22). The addition of these parameters improved the model fit to be comparable to that in the European sample, Satorra-Bentler $\chi^2(162) = 317.54, p < .001$, $CFI = .92$, $TLI = .90$, $RMSEA = .04 [.03 - .05]$.

**Item Response Analyses.** I ran a two factor model loading the AS items on Factor 1 and the PM items on Factor 2 and allowing those factors to correlate. I also included the cross-loadings suggested by the CFA on Factor 2, but I did not drop the poor performing items because IRT can provide insight into why these items perform poorly. I used the software IRTPRO and the default Bock-Aitkin EM estimation method (BAEM; Bock & Aitkin, 1981), which is appropriate for two dimensional models and allows for the estimation of all the available fit statistics provided by the software. The items were estimated using the graded response model for polytomous items (Samejima, 1969; 1997).

This model was not a good fit to the data overall, $AIC = 38245.98$, $M^2 (156) = 538.11, p < .001$, $RMSEA = .06$. However, the individual item fit statistics were generally good with only five items having a significant SS-$\chi^2$ test of model fit (Orlando & Thissen, 2000; 2003). Several items yielded significant tests of marginal fit (LD, Chen & Thissen, 1997). Those tests provide similar information to Lagrange Multiplier tests in CFA, indicating items with significant covariances above those included in the model. The items with the highest LD statistics were those that were given correlated errors in the CFA. Interestingly, the third factor items from the Romanticism facet had test statistics in the moderate range suggesting that they were not causing significant multidimensionality issues.

I ran a second model allowing these items to load together on a third factor: This model required a switch to Metropolis-Hastings Robbins-Monro Monte Carlo (MHRM; Cai, 2010-a;
Cai, 2010-b) estimation for a three factor polytomous model. IRTPRO does not provide the same fit statistics for the model as BAEM estimation, however it does provide the AIC, which in this case was not improved enough to justify the additional complication to the model, $AIC = 38175.52$. The parameters for the two factor model and the individual item fit statistics can be found in Table 4. Unfortunately, IRTPRO does not provide Item or Test Information Function estimates for multidimensional models. These will be reported later as part of the Differential Item Functioning analysis which also required splitting the factors.

The difficulty parameters output from multidimensional models are actually conditional on both dimensions: That is, rather than being able to calculate a point where the unidimensional function crosses $P(\theta_1) = .50$, there is an entire line across $\theta_2$ where $P(\theta_1|\theta_2) = .50$ (Reckase, 1999). This makes it difficult to interpret the difficulty parameters and compare them between items. To ease interpretation, I calculated the difficulty for the factors separately which is equivalent to reporting the multidimensional difficulty the point where $P(\theta_1|\theta_2 = 0) = .50$ ($\theta$s are distributed as standard normal making 0 the mean). So the difficulty parameters are for each subscale holding the other subscale constant at the mean, which is consistent with how I use the measure in substantive analyses. I calculated the IRT parameters for each subscale separately and report the difficulties based on these unidimensional analyses. For AS the difficulty parameters are all quite low, especially Item 9 which is only rated as “Strongly Disagree” by those with a latent trait score of -5.29 standard deviations below the mean. PM on the other hand had quite a bit more variation in the difficulty levels between items and covers a wider range of latent trait levels.

**Differential Item Functioning (DIF).** The multi-group CFA indicated structural differences in the nMyst scale between Chinese heritage and European heritage participants. To see which items were contributing to the difference, I performed a DIF analysis comparing the
item parameters in these two groups. To make it easier to see the differences in the items, I performed the DIF analyses on the individual subscales rather than the full model. IRTPRO uses Wald tests to compare the parameters of the items following Lord (1977) but with accurate item parameter error variance-covariance matrices computed using the Supplemented EM (SEM) algorithm (Cai, 2008). The output provides $\chi^2$ tests for each item overall as well as separate tests for the discrimination and difficulty parameters. Overall DIF test scores and item parameters for the Chinese heritage sample are reported in Table 5.

The AS subscale showed significant DIF in all but two items (Items 5 and 9) on the overall test. Only Item 8 had a significant difference in discrimination, with a higher discrimination parameter for the Chinese heritage participants, $a = 1.44$ compared to $a = 0.73$. Nearly all the items had significant DIF when comparing item difficulty. Inspection of the parameters showed a generally higher trait level necessary at each category level for Chinese heritage participants, suggesting that a Chinese heritage participant would need to be at a higher trait level than a European participant to endorse the same Likert category on the same item. The test information curve (Figure 3) indicates that for both Europeans and Chinese heritage participants the AS subscale is most accurately measuring a fairly low level of the latent trait. In the Chinese heritage sample, the peak information is greater ($I = 9.03$, $SEM = .33$) than the peak information for the Europeans ($I = 6.92$, $SEM = .38$), and the peak information is at a lower trait level ($\theta = -1.40$ compared to $\theta = -.90$).

Most items on the PM subscale showed significant overall DIF, except for Items 25 and 27. Again, most items were not significantly different on the discrimination parameter, only items 14, 15, and 16 and again these items had higher discrimination in the Chinese heritage sample. Most of the DIF was driven by the difficulty parameters; however in this case the item difficulties were generally lower for the Chinese heritage participants meaning that they would
endorse a particular item at a lower level of the latent trait than a European participant. The test information curve (Figure 3) shows that the peak information is much greater for the Chinese heritage participants than the Europeans ($I = 9.35$, $SEM = .33$ compared to $I = 6.62$, $SEM = .39$), but unlike AS the information peak is at roughly the same trait level ($\theta = .10$).

**Religious Conversion.** Participants reported the religion they were raised in as well as their current religious affiliation. A “conversion” variable was calculated based on whether these reported religions were different or not: No change coded as 0 and a conversion coded as 1. Religious affiliations with at least 200 participants were included in the analysis: Atheists, Agnostics, Baptists, Buddhists, Catholics, Fundamentalist Christians, Jews, Mormons, Muslims, and Protestants. By far the most common current belief for converts was Agnosticism, perhaps due to the cultural bias against Atheists in North America (Edgell, Gerteis, & Hartmann, 2006). Table 6 shows the number of participants in each current by childhood religion to give a more thorough picture of the number and type of conversion experiences represented in this sample.

Before examining the conversions, I ran an ANOVA to test my hypothesis that religious affiliation would relate to differences in nMyst. I used only the non-converts because they would be most typical of the particular religious belief, leaving a sample of $N = 6393$. The ANOVA was significant for both AS and PM, $F (10, 6383) = 253.26, p < .001$ and $F (10, 6383) = 53.87, p < .001$. Atheists were lowest on both AS and PM, though Baptists and Mormons were not significantly higher on PM. Baptists and Fundamentalist Christians were highest on AS, all others were significantly lower. Buddhists were highest on PM, significantly higher than Agnostics and Catholics who did not differ at all. These results are shown in Figures 4, AS, and 5, PM, as the “Non-Convert” bars.

Four ANOVA models were run (for ease of interpretation) for current and childhood religious affiliation with AS and PM as dependent variables. People who converted from their
childhood religion were significantly different on AS from those who were raised in the same religious tradition, but did not later change their beliefs, \( F(1, 13652) = 146.29, p < .001 \). To rule out the possibility that it is simply a loss of religious belief based on changing scientific attitudes, I removed all those who had converted to Agnosticism or Atheism. The main effect of conversion was marginally significant, \( F(1, 7407) = 3.42, p = .06 \). However, the interaction between conversion and religion was significant, \( F(7, 7407) = 7.34, p < .001 \). The fact that the direction of the conversion effect varies between different religions explains the lack of a significant main effect.

For current belief using all religions, there was no main effect of conversion, \( F(1, 12484) = .02, p = .90 \). However, there was an interaction between religion and conversion, \( F(1, 12484) = 4.98, p < .001 \), suggesting that some of the religions did show a difference between converts and those raised with those beliefs. The simple effects showed that there were significant differences between Atheists, Agnostics, Buddhists, and Fundamentalist Christians. However, all of these differences showed rather small effect sizes with only Atheists showing a mean difference greater than .20. In each case those who converted held more “extreme” views than those who held the beliefs their whole lives, i.e. convert Atheists were even lower on AS, and convert Fundamentalists were even higher. Figure 4 shows the mean comparisons for all religious affiliations on AS with childhood and current religion comparisons on the same graph.

Pro-Magic showed a similar pattern to Anti-Science. The first model again compared those who were raised in a particular religion and subsequently converted (converts-from) with those who maintained their beliefs (non-converts). There was a main effect of conversion, \( F(1, 13632) = 18.20, p < .001 \), suggesting that those who converted are significantly different from those who did not change their beliefs. In general, those who converted seem to have a more average level of PM than those who stayed in the particular religion. The conversion main effect
was not significant when comparing converts to non-converts on their current religion, $F (1, 12484) = 1.63, p = .20$. Again, the interaction was significant suggesting that some of the individual religions did show significant differences, $F (9, 12484) = 2.99, p < .001$. Significant differences were found for Atheists, Buddhists, Mormons, and Protestants. Only Mormons had a mean difference greater than .20, with those who converted being higher on PM than non-converts. Figure 5 shows the pattern of results for each religion comparing converts and non-converts in both childhood and current religions.

**Demographics.** In contrast to the student sample in Study 1, this sample did show significant gender differences. Women were higher on both AS and PM, $M_{\text{diff}} = -0.42, t (12,437) = -26.96, p < .001$, and $M_{\text{diff}} = -0.25, t (12,437) = -17.74, p < .001$, respectively.

**Ethnicity.** 12 possible ethnic backgrounds were included as options in the questionnaire and all had a significant number of participants: Native American or Alaskan Native, Black or African American, Chinese, Filipino, Hispanic, Indian, Japanese, Korean, Middle Eastern, Pacific Islander, Vietnamese, White or European heritage, and Other Asian. The sample was fairly mixed in whether people were located in their home country or in the US (though nearly all Chinese participants were located outside China, likely due to restrictions on internet use within the country). I ran an ANOVA comparing the means of AS and PM for all 12 groups. Both AS and PM were significant, $F (12, 14406) = 6.39, p < .001$, and $F (12, 14406) = 8.48, p < .001$, respectively. The mean ranges were surprisingly small, from 3.10 to 3.52 for AS and 2.46 to 2.89 for PM. Individual group comparisons were significant even when differences were very small due to the large number of participants. I only report differences greater than .15. Comparing other groups to the European heritage group on AS, only the Native American and African American groups differed by more than .15 and were both higher on AS. The Japanese and Indian participants were the lowest on AS. For PM, Native American and Vietnamese
participants were higher than European heritage participants, and Koreans were significantly lower. The Korean and Indian participants were lowest on PM. Figure 6 shows the overall ANOVA results.

Importantly, the internal consistency reliability of the nMyst scale was acceptable, and indeed, very similar across groups. AS reliability was greater than .80 in all but four groups, the lowest being Filipino participants with $\alpha = .74$ and the highest European heritages with an $\alpha = .86$. PM was somewhat lower overall, with all but one group falling between .72 and .80. The Vietnamese group reliability was much lower at .66. Native American participants had the highest reliability with $\alpha = .80$. All the exact reliabilities can be found in Table 7.

**Education.** Highest level of education reported was recoded into numerical categories ordered by the number of years of education generally associated with that level. The sample ranged from “Some High School” ($N = 731$) to “PhD/Post Doctoral” ($N = 292$). Most participants had some college, but had not completed a Bachelor’s degree ($N = 6,284$). Treating the recoded education attainment as a continuous variable, it was uncorrelated with either AS or PM, $r = -.07$ and $r = -.09$, respectively. However, an ANOVA treating it as categorical did show significant differences for both AS and PM, $F (5, 14595) = 24.88, p < .001$ and $F (5, 14595) = 24.56, p < .001$, respectively. The plot suggests that the most significant decrease in nMyst occurs at the highest education level and the differences until that point are quite small between adjacent categories, e.g. $M_{diff} < .15$ (Figure 7).

Education has often been hypothesized to cause rejection of religious belief. Education could potentially explain why lower AS was related to change to either Atheism or Agnosticism. To test this relationship, I used a chi-squared test on the cross-tab of education level and those who converted to Atheism and Agnosticism ($N = 4,215$). The test was not significant,
suggesting that those at a particular level of education were no more likely than any other education level to reject religion, $\chi^2 (5) = 6.08, p = .30$.

**Personality.** All correlations greater than .04 were significant due to the sample size, only the effect size (magnitude of the correlation) was considered in reporting results in this section. Participants of all ethnicities were included in these results. AS was correlated with Agreeableness, partial $r = .26$ controlling for PM. PM was correlated with Neuroticism, partial $r = .16$, and Openness to Experience, partial $r = .15$, also controlling for AS.

**Discussion**

**Confirmatory Factor Analysis.** The final item set after dropping items in this study can be found in Appendix C. I started with a three factor model reflecting the EFA results from the Study 1. Model 1 did not fit the data so I continued my exploratory analysis by running LMTESTs to find parameters that could be released to improve fit. I began with cross loadings and ended up releasing 5, specifically Items 12 and 22 cross-loaded from Factor 1 to Factor 2 and the secondary loadings were positive. I also tested two factor models to see if an unadjusted two factor model or two factor with cross-loadings would be a better place to start my exploratory model adjustment, but all two-factor models had inferior fit to the three factor versions.

Item 22 is an intuition item and these items had loadings on both factors in the EFA and are part of the more general theory of nMyst. Item 12 is a First Person Perspective item, but it also references spiritual visions which are often part of the more new age spiritual experiences associated with PM. Items 16 and 17 are Factor 2 Animism items which cross-loaded on Factor 1, but their secondary loadings are actually negative. The final cross-loading is a Factor 2/3 item which loads significantly on both. The loading is perfectly consistent with my theory that Romanticism should be part of the PM factor. The error correlations released are all within facet
so it is natural that they would be correlated with each other above their correlations with the overall construct.

All the cross-loadings and all but one of the correlated errors were still significant in the confirmation dataset. Their persistence across datasets and simple explanations for why they would be significant suggest that these additional parameters are actually characteristics of the data rather than capitalizing on chance associations.

The question still remains whether the two factor or three factor solution is best. Although neither showed an impressive fit, there was little difference between the unadjusted 2 and 3 factor models. The increased adjustments lead to an acceptable fit using the three factor model, but neither was much better without significant adjustments. In actual use, all the error covariances within factor will be accounted for by creating a composite and the cross-loadings are accounted for using regression to control for the other factor. Basing the composites for the subscales on the two factor solution does ignore some difference between the Romanticism facet and the other Pro-Magic facets, but the theory is based on their commonality, which is a large portion of their variance given the high correlation. Further validity evidence will also support the consistent interpretation of the two factor solution.

**Multigroup CFA.** The disadvantage of the three factor structure was most evident in the Chinese heritage sample. The MGCFA found serious differences in the factor structure between the two groups. The two factor model was a better fit in the Chinese heritage sample. It was encouraging to see that many of the suggested additional parameters were the same between the two samples, though a few additional error covariances were suggested that were not within facet and item 25 was actually so low loading that it clearly should be dropped from the scale. It was also the lowest loading item in the European sample as well ($\lambda = .35$). This analysis does suggest
that if I continue to use the scale as two dimensional it should actually perform equally well, if not better, with Asian participants.

**Item Response Theory.** These results shed some light on the interpretation of scores on the Anti-Science subscale. The difficulty parameters suggest that the items on the AS subscale are very “easy”: That is, one does not require very strong anti-science attitudes to endorse these items. The low item difficulty would explain the fairly high average scores on the scale (item mean of 3.40 in the current sample). The item information metric shows that the peak measurement accuracy is nearly 1.5 standard deviations below the mean on the latent trait. Adjustments could be made to the items in the future to make the wording slightly more extreme to more accurately capture higher levels of the latent trait. Generally, the items show reasonable discrimination, not unexpected given the CFA loadings.

Compared to AS, the PM subscale showed more variation in difficulty and the peak information was actually near the mean latent trait meaning that it is maximizing measurement accuracy at the most common score level. However, the discrimination parameters are not as strong as those for the AS subscale. The items that were dropped from the CFA analysis, Items 29 and 19, had acceptable discrimination parameters though Item 19 was one of the lowest, \( a = 1.04 \) and \( a = 0.78 \) respectively. Item 29 had fairly low difficulty compared to other items which could explain its low loading if it is being answer differently from other items on the scale. This analysis also supported dropping Item 25 which had a very low discrimination parameter.

**Differential Item Functioning.** Although the rate was low, most of the instances of DIF between the European heritage and Chinese heritage subsamples showed better item functioning in the Chinese heritage sample. The difficulties of the AS items were closer to the latent trait mean. Total scale information was better for both subscales meaning that measurement in the in the Chinese heritage group is actually more accurate for the whole scale.
Religious Conversion. Comparing different religions gives some more insight into how the two strategies of nMyst interact with each other. Some religions are above the mean on both AS and PM, such as Catholicism and Islam. Others employ only the AS strategy, in this case all Christian denominations other than Catholic. The non-Christian religions tend to be above the mean only on PM, e.g. Hinduism, Judaism, Buddhism, as well as Agnostics. Atheists were the only group who were below the mean on both strategies.

These patterns fit with particular aspects of the different religions. Mono-theistic religions tend to employ the AS strategy, either exclusively or in conjunction with PM. The exception is Judaism, which historically has been paired with a strong cultural emphasis on education and science knowledge. Catholicism uses both strategies; this seems to be consistent with the more magical elements of Catholic dogma, such as transubstantiation and worship of saints (Mead & Hill, 2001). Religions with less focus on a central deity figure, such as Hinduism, or Buddhism, are less threatened by scientific explanations and rely more on the PM strategy.

The results generally supported my hypothesis that people who choose to change from a particular religion have nMyst scores that are different from the people who choose to stay in the same religion. This finding held even when not including “conversions” to Atheism or Agnosticism. Though generally people were not significantly different on nMyst from people who grew up in their new chosen religion, there were some consistent differences. Converts tended to be more extreme that the non-converts in their new religions, i.e. Fundamentalist Christians being even higher on AS. I also found a difference between Western Buddhist converts vs. traditional Buddhists in that the traditional Buddhists had more average nMyst scores (slightly higher on both subscales than converts). The difference between Western and Eastern Buddhist is possibly due to the cultural institutional aspect of Buddhism in Asia making
it more like a traditional religion. In the Western world it is treated as an alternative to traditional religious beliefs and attracts people more similar to Agnostics (Coleman, 2002).

These results suggest that a longitudinal study could be worthwhile since there is some initial cross-sectional evidence. A longitudinal study would directly address whether the difference in nMyst precedes the conversion or is the result of learning the beliefs of the new religion. Interestingly, nearly half of the participants in this sample report a current religion that was different from the religion they were raised with suggesting that conversions are not terribly rare. However, the majority of conversions were to agnosticism so a large sample would still be necessary to capture enough conversions not based primarily on anti-science beliefs.

**Demographics.** Unlike the student sample, the community sample did show a gender difference with women being higher than men on Anti-Science. The difference between samples could be due to the general tendency for women to be less interested in math and science in school, which was washed out by the lack of education variation in the student sample. It could also be explained by the fact that women tend to be more strongly religious than men (Pew Research Center, 2007).

Overall, the ethnicity results indicate a reasonable degree of cross-cultural validity. The nMyst scale seems to tap a coherent construct across a large variety of ethnic/cultural groups. Europeans tend to be fairly low on nMyst compared to other groups, but certainly not the lowest. Indians were lower on both AS and PM, which could be a result of the surge in science and technology education in India, especially among people who would have a Facebook account. Overall, however, the differences in group means are not substantial: The small effect sizes suggest that a similar rate of mysterious beliefs emerges in different cultures despite obvious differences in worldviews.
The reliabilities show that both subscales are similarly coherent across cultures. This finding is encouraging because it suggests that the scale can be used and interpreted in cross-cultural studies. It can also be used in multicultural samples, such as the students at UBC, without needing to separate cultural groups. The consistency across cultures also supports the conception of nMyst as a need and not simply a preference. The basis for mysterious beliefs is that they evolve naturally out of universal explanatory tendencies that underlie all religious belief. Any theory based on evolution requires that the construct be generalizable to different cultures.

The association of nMyst with education was surprisingly weak. It appears that education has only a small relationship with nMyst until the very highest levels. This sample was sufficiently diverse in their education experiences. Though most participants had some college experience, there were a large number of participants who had not attended college and many who had graduate or professional degrees. Either causal direction could be argued, that lower nMyst leads to pursuing more education, or that education causes lower nMyst. However, with the small effects of education in the more common range it suggests that even if education is a cause of lowering nMyst it is not enough to account for the difference found in the conversion results. I would conclude that simply acquiring more education is not sufficient, in itself, to cause a person to reject religion -- unless they start with an already low level of AS or PM.

**Personality.** The Big Five correlations found in this dataset were similar to those found in Study 1. Supporting a trait interpretation of nMyst, the correlations between AS and Agreeableness, and PM and Neuroticism and Openness have now been replicated in both student and community samples so it is reasonable to assume that they are reliable.
Chapter 4: Political Attitudes

Study 3 examined the association of nMyst with a number of measures related to political attitudes. This investigation is predicated on the longstanding evidence that personal beliefs play a significant role in broader worldviews.

In particular, there is a long tradition of research linking political conservatism with anti-science attitudes (Eysenck & Wilson, 1978; Rokeach, 1973). More recently, the link between motivated cognition and political conservatism has played a role in System Justification Theory (Jost & Banaji, 1994). Feelings of threat or fear stimulate ideological motives that influence political conservatism (Matthew, Levin, & Sidanis, 2009). These ideological effects have been shown in response to direct fear of attack (e.g. the terrorist attacks on September 11th), fear of change to the social order, or fear of death (Bonanno & Jost, 2006; Conover & Feldman, 1981). This defense of the status quo is especially pronounced in the constellation of variables linked to conservatism: Authoritarianism (Adorno, et al., 1950), Belief in a Just World (Lerner, 1980), Social Dominance Orientation (Pratto, Sidanis, Stallworth, & Malle, 1994), and Need for Structure (Kruglanski & Webster, 1996).

Anti-Science attitudes (AS) also appear to reflect that same motivation to reject new information that threatens their beliefs. Given that information from scientific quarters tends to continually change, the subsequent uneasiness can be dealt with by developing a generally skeptical attitude toward science. Instead, conservative ideologies provide a more firm and consistent worldview. Therefore I hypothesize that AS should be related to the constellation of variables associated with conservatism. I will discuss these variable in greater detail below.

Belief in a Just World. This construct taps the tendency to assume that everything that happens to a person is deserved (Lerner, 1980). All successes and failures are the direct results of a person’s actions. Belief in a just world is a comforting belief because it provides people
with the sense that if they do what they are told then they will be rewarded and nothing bad can happen. People who experience misfortune brought it on themselves by their behavior. This belief underlies victim blaming for crimes such as rape (e.g. “She shouldn’t have been in that neighborhood alone”) and can be used to justify ending social welfare programs (e.g. “They’re all just lazy”).

**Right Wing Authoritarianism** (Adorno, et al., 1950, Altemeyer, 1981; 1998). This classic construct concerns the belief that established authority and tradition is always right and should be followed. It is also related to a desire to censor new ideas that could promote rebellion. It overlaps with just world belief in that it assumes that the current authority exists due to its own merit and competence and should not be questioned. The Right Wing Authoritarianism (RWA) scale was developed by Altemeyer (1981).

**Views of morality.** The Moral Foundations Questionnaire (Graham, Haidt, & Nosek, 2009) carves out the specific moral issues that form the basis of larger worldview tendencies. One of the key points of the Moral Foundations theory is that the clash between liberal and conservative worldviews derives, not from the fact that one view is morally superior, but from the fact that the two ideologies resonate to different values (e.g., Haidt, 2008; Graham, Haidt, & Nosek, 2009). Endorsing a foundation means agreeing that issues in that topic are morally relevant. In early work on these foundations, Haidt looked specifically for scenarios where liberals and conservatives would differ on whether an act was morally wrong. These moral judgments were primarily based on emotional responses like anger or disgust.

Through testing a number of these types of scenario Haidt settled on five general foundations that are present in numerous cultures: Harm/Care, Fairness/Reciprocity, In-group/Loyalty, Respect for Authority, and Purity/Sanctity. Some of these are easily understood: Harm to others is a very common basis for moral judgments. Fairness is also a common
concern. Research on dictator games in economics, where one participant allocates some portion of a reward between themselves and another participant, often shows a tendency to split evenly even when there is no possible reciprocity (Fehr & Schmidt, 2006). Liberals primarily consider Harm/care and Fairness/reciprocity as moral concerns: Those foundations are labeled individuating, due to their focus on the rights of the individual. Along with those two, conservatives also endorse In-group loyalty, Authority/respect, and Purity/sanctity.

In-group loyalty is the belief that concern for the in-group over outsiders is actually a moral imperative. As an example, people high in this foundation would support harsh punishment for treason against one’s country. Respect for authority refers to the sense that it is morally wrong to disobey an authority figure (e.g., the government or one’s parents). It is also related to authoritarianism because the authority figure is seen as morally infallible and deserving of their position of power. The purity foundation is often related to religious moral laws about hygiene or food consumption. Examples include Halal or Kosher food preparation rules as well as rules prohibiting sex outside of marriage or homosexuality. Those foundations are labeled binding because they emphasize the importance of maintaining the social order (Haidt, 2008; Graham, et al., 2011).

Most people, regardless of their political views, agree that harm and fairness are moral concerns. The other three show some disagreement, liberals tend to prioritize the individuating foundations and generally don’t consider the binding foundations to be moral issues. Conservatives tend to endorse all five foundations giving almost equal weight to harm and fairness as they do to respect for authority in a moral dilemma. Because nMyst is tied in with spiritual beliefs, I expect that both subscales will be in some way related to all of the moral foundations. AS should be more related than PM to the conservative foundations. However, PM may still be related to Purity/sanctity as a result of seeing more things as sacred.
**Free Will Belief.** Believing in free will fits the definition of mysterious beliefs as being teleological rather than mechanistic (Carey & Paulhus, in press; Paulhus & Carey, 2011). That belief requires the assumption that there is something special about human agency and that it comes, not from the brain or environment, but some other aspect of being human. Free will belief is consistent with belief in the soul, or mind-body dualism, which is also a teleological belief. I expect that both PM and AS will correlate positively with belief in free will because the latter is related to conservative beliefs (especially religiosity) as well as a more general orientation towards teleological explanations. The FAD-Plus (Paulhus & Carey, 2011) scale also measures belief in fatalistic determinism, the belief that events in our lives and even our actions are result of outside forces such as fate. Belief in fate fits very well in the nMyst worldview because it implies a universal reason for anything and everything based on a mysterious agentic force. It is also a type of determinism that does not contradict anti-scientific sentiments because it is not based on scientific causes. For this reason I expect that the fatalistic determinism will be positively correlated with both aspects of nMyst.

The FAD-Plus scale also includes two other subscales for which I made no a priori hypotheses. The Scientific Determinism subscale addresses determinism from a more rational standpoint of scientific cause (e.g. the effects of biology or social structure on behavior). The Unpredictability subscale taps the extent to which one believes that things that happen have no causes. The four beliefs are relatively uncorrelated (Paulhus & Carey, 2011).

**Method**

Participants were recruited using the website Mechanical Turk ($N = 360$, 58% female, age range 18 – 70, mean age 33.5, 50% European, 22% East Asian). Respondents were restricted to be from North America to avoid issues of English language proficiency and ethical payments. They were given the 32 item version of the nMyst scale, though only the final 20
items were used in constructing the subscale composites. The BFI, Moral Foundations Questionnaire (Graham, Haidt, & Nosek, 2009), Belief in a Just World (Lerner, 1980), and Right Wing Authoritarianism (Zakrisson, 2005) were also included. Again religious belief and religiosity were collected along with standard demographics.

**Measures**

**FAD-Plus.** Developed by Paulhus and Carey (2011), this instrument consists of 27 items in 5-point Likert format. The four subscales are Free Will (7 items), Scientific Determinism (7 items), Fatalistic Determinism (5 items) and Unpredictability (7 items). In the present samples, the alpha reliabilities ranged from acceptable to solid: Free Will (.70), Scientific Determinism (.69), Fatalistic Determinism (.82), and Unpredictability (.72).

**Authoritarianism.** Participants were given a 15-item version of Altemeyer’s 23-item (1981) RWA scale. Using more modern wording, it was developed by Zakrisson (2005). Despite the shorter length, scale showed a solid alpha reliability in this sample, α = .83.

**Belief in a just world.** This concept was measured with Lerner’s (1980) 12-item individual difference measure. Only the six-item subscale tapping belief in a just world for others (JWB-others) was analyzed here. In this sample, alpha reliability was .90.

**Moral foundations.** The Moral Foundations Questionnaire (MFQ) is a 30-item 6-point Likert format instrument with growing support (Graham, Haidt, & Nosek, 2009). As is common with this measure, alpha reliabilities for the subscales were on the modest side: Harm, α = .56, Fairness, α = .53, In-group, α = .61, Authority, α = .60, and Purity, α = .72.

**Results**

The nMyst subscales were assembled using the item set from Study 2. Reliabilities were comparable to previous samples (AS, α = .83; PM, α = .78) and the intercorrelation was slightly higher, r = .45.
Demographics. Women scored higher than men on AS, $M_{\text{diff}} = -0.19$, $t = -2.55$, $p < .01$. PM had no gender difference, $M_{\text{diff}} = -0.03$. Age was negatively correlated with PM, $r = -0.18$, $p < .001$, but was not related to AS. There was no change in the correlations with age when controlling for both religiosity and the other subscale.

Political Attitudes. Regression was used to control for relationships among all the political attitude measures as well as religiosity and the other nMyst subscale. All MFQ subscales, Right Wing Authoritarianism, and Belief in a Just World were entered together with the nMyst subscale as the dependent variable. It was assumed that most of these variables would be positively correlated, so a regression analysis was chosen to avoid the numerous third variable problems and to control Type 1 error.

AS was positively related to RWA, $\beta = .17$, $t (9) = 3.08$, $p < .01$. Belief in a Just World was not related either subscale of nMyst. AS was significantly related to moral concerns about Harm, $\beta = .13$, $t (9) = 2.53$, $p < .05$. It was negatively related to the binding foundation of In-group/Loyalty, $\beta = -.13$, $t (9) = -2.51$, $p < .05$, and positively related to Purity, $\beta = .32$, $t (9) = 4.81$, $p < .001$. Interestingly, there was an unexpected correlation between PM and In-group Loyalty, $\beta = .28$, $t (9) = 4.54$, $p < .001$. PM was also negatively related to the Authority foundation, $\beta = -.15$, $t (9) = -2.12$, $p < .05$.

Free Will. PM and AS were correlated with the Free Will subscales controlling for religiosity and the other subscale. AS was positively correlated with belief in free will, partial $r = .19$, $p < .001$, and fatalistic determinism, partial $r = .23$, $p < .001$. AS was negatively correlated with belief in scientific determinism, partial $r = -.15$, $p < .01$. PM was negatively correlated with belief in free will, partial $r = -.13$, $p < .05$. PM was positively correlated with both forms of determinism, fatalistic, partial $r = .35$, $p < .001$, and scientific, partial $r = .30$, $p < .001$, as well as Unpredictability, partial $r = .20$, $p < .001$. 
Discussion

Demographics. Consistent with the Facebook sample, I found a gender difference with women scoring higher than men on Anti-Science. Inconsistent with the Facebook data, women in the current sample women were not higher than men on PM. This discrepancy could be the results of the greater diversity of the Facebook sample on several different variables.

The negative correlation between PM and age is consistent with research on paranormal beliefs (Emmons & Sobal, 1981). Younger people are more likely to still be deciding what they believe and are more likely to endorse magical or paranormal beliefs. I would have expected AS to be positively correlated with age due to its relationship with religiosity which generally increases with age. It may be that an AS worldview is more stable across time independent of the association with religiosity.

Political Attitudes. Given the theoretical overlap among the current measures of political attitudes, I felt it important to analyze them together to ensure a correct situating of nMyst in the nomological net. As hypothesized, AS was positively related to Right Wing Authoritarianism even when controlling for religiosity. This relationship probably accounts for the lack of association between AS and other conservative attitudes such as Belief in a Just World. PM was not related to either of these, as was hypothesized.

The results of the Moral Foundations Questionnaire were more complex than suggested in my hypothesis that only binding moral foundations would be related to nMyst. Instead, AS showed a small positive correlation with the Harm foundation: However, there is nothing in the AS attitude that suggests that they should not prioritize harm as a moral concern -- other than a general politically conservative ideology. But conservative ideology was already accounted for by including RWA in the regression model.
These speculations are necessarily tentative; given the small size of the correlations and the fact that they were unexpected. It’s possible that they are artifacts of including so many predictors in the model. The one correlation that was predicted, between Purity and AS, was actually of moderate effect size. Those who use the anti-science strategy hold all of their beliefs as totally sacred and that they must be protected so it is understandable that they would see purity violations, which are often tied to the sources of their religious worldview, as important moral concerns.

The correlations with PM were also unexpected. I hypothesized that PM would likely relate to the Purity foundation, instead the largest correlation was between PM and In-group/Loyalty even after controlling for RWA. The negative correlation between PM and Authority does make sense because in some cases these beliefs are a rejection of organized religions and their hierarchical structure.

**Free Will.** I found the hypothesized positive association between AS and belief in Free Will. Free Will is itself a teleological view of human behavior related to mind-body dualism. It is also often being threatened or challenged directly by neuroscience and philosophy. The negative correlation of PM with belief in free will was unexpected. Instead, PM correlated positively with all other subscales -- fatalistic determinism, scientific determinism, and unpredictability. Most PM beliefs entail content that places agency in outside actors who are responsible for things that happen, e.g. ghosts or the alignment of the planets: Hence a positive correlation with fatalistic determinism makes perfect sense.

These results may also provide an explanation for the lack of correlation between PM and Free Will. Religious teachings, specifically Protestant Christianity, account for the coexistence of free will and the idea that God has already decided your future (Myers, 2008). Without religious explanations for their coexistence, fatalistic determinism is logically inconsistent with
free will. In fact, the philosophical position of Hard Determinism stipulates that these two
concepts are mutually exclusive and cannot be true of the same universe. Perhaps those high on
PM, and also not religious, take the fatalism belief to its logical conclusion leading to a negative
correlation with belief in free will.
Chapter 5: Discriminant Validity and Temporal Stability

Thus far I have only addressed the internal consistency reliability of the nMyst scale. Although the alpha reliability statistic is the traditional measure of internal consistency, it only captures the instrument’s reliability at a moment in time. It does not capture temporal stability of the instrument. Therefore, it has become standard practice to report test-retest reliability. Indeed, some authorities on the subject have argued that test-retest is a more important statistic than internal consistency (McCrae et al., 2011).

The test-retest reliability value indicates the consistency of an individual’s responses when taking a scale multiple times. It directly taps the stability of the concept by showing whether people interpret the items in the same way every time. In this study, participants completed the nMyst scale on two separate occasions approximately three months apart.

In Studies 1-3, I have focused on the convergent validity evidence, trying to determine which conceptually related items were in fact correlated with the nMyst scale. In this study, I will address another aspect of the nomological net, namely, discriminant validity. A few variables have been suggested as possibly being entirely redundant with nMyst conceptually. It is important to show that nMyst is not just reinventing the wheel by duplicating an existing variable. Schizotypal personality disorder (SPD), Absorption (Tellegen & Atkinson, 1974), and Mysticism (Hood, 1975), are all theoretically related to the PM subscale and have been brought to my attention as variables that should be distinguished from nMyst.

Recall that PM involves a tendency to see every day experiences as spiritually meaningful, as well as a tendency to believe in less commonly accepted supernatural explanations. The DSM-IV-TR criteria for Schizotypal Personality Disorder include a number of symptoms related to paranormal belief: magical thinking and odd beliefs, unusual perceptual experiences, ideas of references (the belief that things in the world are happening specifically
because of or for you), and paranoid ideation (DSM-IV-TR; 2000). Magical thinking and odd beliefs are an explicit aspect of the Pro-Magic strategy, and the other symptoms listed can lead to experiences of awe or a spiritual interpretation of mundane experiences. There are also some aspects of SPD that would not have any relation to PM: Constricted affect, lack of close friends, or social anxiety. They are more clearly impairing and I would not expect them to be part of nMyst. I am certainly not suggesting that using the PM belief strategy is the results of having a personality disorder, but that in the normal population there are many similarities in the concepts. I hypothesized that the magical/paranormal aspects of SPD would be highly correlated with PM in a normal population, but things like paranoia would be less correlated.

The concept of Absorption came about from researchers trying to find a personality trait that predicts susceptibility to hypnotism (Tellegen & Atkinson, 1974). They interpret absorption as a particular state of attention to an experience or object. This state is associated with a sense of hyper-reality of the focus of attention, imperviousness to outside distractions and a sense of alteration to the self. The authors’ Absorption scale measures a tendency or desire to have such experiences. The scale is, as intended, related to hypnotizability as well as Big Five openness to experience (Glisky et al., 1991). Certain items contain similar content to the Romanticism facet of PM. Also the idea of having an appetitive motivation to have these experiences is similar to the motivation to have new spiritual experiences and beliefs that characterizes the PM strategy.

Mysticism is the experiential component of intense spiritual and religious experiences. The construct is built around eight criteria identified as being a part of intense mystical or religious experiences (Stace, 1960): Ego quality, unifying quality, inner subjective quality, temporal/spatial quality, noetic quality, ineffability, positive affect, and religious quality. Many of these qualities have similar meanings or interpretations to facets included in both subscales of nMyst. Ego quality refers to a loss of the sense of self as a unique entity and absorption into
something greater, unifying quality is the sense that everything in the universe is part of a single “whole”. Inner subjective quality is the sense that all things have an inner consciousness, which is similar to the concept of animism. Temporal/spatial quality is the sense that the experience is somehow disconnected in time or space or that during the experience the ideas of time and space become meaningless. Noetic quality is the sense that experience is valid as a source of knowledge, similar to First Person Experience. Ineffability is the indescribable quality of the experience. Positive affect and religious quality are perhaps the most straightforward, these experiences are expected to provide a sense of profound happiness and the religious quality allows for experiences to be described in specifically religious terms such as “sacred” or “holy”. The phrasing of the items asks participants to rate whether they have had experiences with these various qualities.

The Mysticism scale has been found to break into religious specific and generally spiritual factors similar to the two factor structure of nMyst (Hood, 1975). I expect that the general aspect of Mysticism would be positively correlated with PM and the religious aspect would be correlated with both aspects, but more strongly with AS.

Method

Participants were recruited through the UBC Human Subject Pool (N = 100). This study was only available to participants if they had completed a 17 item short measure of nMyst as part of the prescreening questionnaire. The prescreening only allows one minute of items per lab so the scale was shortened as much as possible. The questionnaire is open to all HSP participants and is done voluntarily at the beginning of the term.

Approximately three months later, they were given the full 20 item nMyst scale used in Study 3, as well as the Ideas of References, Magical Thinking and Paranormal Belief, and Unusual Perceptual Experiences subscales of the Schizotypal Personality Questionnaire (Raine,
1991), the Tellegen Absorption Scale (1974), and Hood’s Mysticism Scale (1975). Pre-
screening and study data were linked using the student’s HSP ID created from a combination of 
their birthdays and their parent’s birth months.

**Measures**

**Schizotypal Personality Questionnaire.** The full SPQ has nine subscales and 74 items 
based on the DSM-III-R criteria for Schizotypal Personality Disorder (Raine, 1991). Most of 
these subscales include symptoms that have no theoretical relation to nMyst. I chose 25 items 
from the Ideas of References ($\alpha = .86$), Magical Thinking and Paranormal Belief ($\alpha = .85$), and 
Unusual Perceptual Experiences ($\alpha = .77$) subscales. The exact items are listed in Appendix D.

**Tellegen Absorption Scale.** The Absorption scale has 17 items. The reliability in this 
sample was strong, $\alpha = .86$.

**Mysticism.** The full Mysticism scale has 32 items addressing the eight criteria. These 
items were found to form two factors: General Mysticism, and Religious Mysticism. Due to the 
time constraints of the study, I chose the 22 highest loading items. These items can be found in 
Appendix D. The general subscale reliability was $\alpha = .83$. The religious subscale reliability was 
not as strong, $\alpha = .73$.

**Results**

Out of 100 participants, 71 correctly reported their HSP ID in the study allowing us to 
link their data. The demographics were similar to previous student samples, 68% Female with a 
mean age of 20.5, and 64% of Asian heritage. We included all ethnicities given that Study 4 
confirmed the two factor structure and comparable item functioning in Asian participants.

**Test-Retest Reliability.** The correlation between participants’ initial nMyst scores in the 
pre-screening and their nMyst scores from the study were used as the measure of test-retest 
reliability. The Anti-Science subscale had the same nine items in both the prescreening (Time 1)
and the questionnaire (Time 2). Pro-Magic had only eight items at Time 1 and 11 were included at Time 2. The test-retest reliability for the AS subscale was \( r = .82 \), which was almost identical to its alpha reliability at Time 2, \( \alpha = .84 \). The test-retest reliability for PM was \( r = .70 \) when comparing the full subscale at Time 2 and \( r = .76 \) when comparing only the items included at both times, which was less than the alpha reliability at Time 2, \( \alpha = .81 \). Both test-retest values were highly significant, \( p < .001 \).

**Discriminant Validity.** Correlations among the total scores for SPD, Absorption, and Mysticism were all fairly large \( (r > .4) \) so I chose to analyze the data using regressions to control for the overlap between constructs as I did in Study 3. The raw correlation between the SPQ and the PM subscale was very high, \( r = .72 \). Also the correlation with the AS subscale was significant, \( r = .35, p < .001 \). Mysticism was positively correlated with both AS and PM, \( r = .29, p < .001 \) and \( r = .23, p < .001 \), respectively. However, it was also correlated with religiosity, \( r = .29, p < .001 \). Absorption was positively correlated with AS, \( r = .21, p < .05 \), and PM, \( r = .56, p < .001 \). For the regression analysis, I split the SPQ and Mysticism into their separate subscales.

First, I ran a regression with AS as the dependent variable entering PM, the total SPQ, religious mysticism, general mysticism, absorption, and religiosity. Only religious mysticism remained positively related to AS, \( \beta = .42, p < .001 \). General mysticism and absorption actually became negatively related to AS, \( \beta = -.22, p < .05 \) for both effects. There was no significant relationship with the SPQ, \( \beta = .05 \).

For PM, I started by testing the relationship between PM and the SPQ subscales while controlling for the other subscales. Only the Magical Thinking subscale was uniquely related to PM, \( \beta = .48, p < .001 \). I then ran a regression model with PM as the dependent variable controlling for AS and religiosity with the mysticism subscales, absorption and the magical thinking subscale of the SPQ. Magical thinking remained significant, \( \beta = .53, p < .001 \), as did
Absorption, $\beta = .19, \ p < .05$. General mysticism was no longer significantly related with PM, $\beta = .08$, and religious mysticism was significantly negatively related, $\beta = -.30, \ p < .005$.

**Discussion**

**Test-Retest Reliability.** The test-retest reliability values were strong and comparable to the internal consistency reliability estimates. In other words, participants were generally consistent in how they answered the questions at Time 1 and Time 2 (at least in rank order sense). Although the exact date at which each participant took the pre-test is not known, it ranged between 9 and 13 weeks prior to Time 2. That delay is long enough to rule out the possibility that participants simply repeated their initial responses to be consistent (John & Soto, 2007).

**Discriminant Validity.** As expected, scores on SPD, Absorption, and Mysticism were all intercorrelated. Together, they allowed me to fill in the Pro-Magic portion of nMyst’s nomological net. I included only the most related subscales of the SPQ due to time constraints on the questionnaire. I felt it was unlikely that nMyst would be related to the negative SPD symptoms or more severe forms of paranoia that were unlikely to appear in a non-clinical population. Of those three subscales, only Magical Thinking/Paranormal Belief ended up being positively correlated with the Pro-Magic subscale: Indeed, these beliefs are part of the primary content of PM. All three subscales of the SPQ were unrelated to AS, as predicted, once PM was controlled for in the regression analysis.

The regression results showed that AS was only related to the religious experience subscale of Mysticism and was negatively related to general mysticism and absorption. It makes sense that AS would only relate to mystical experiences in a religious framework. The AS strategy relies on the avoidance of competition for one’s chosen belief or worldview and having quasi-religious experiences outside that framework could be seen as competition.
The Pro-Magic subscale was significantly related to Absorption and Magical Thinking, but not General Mysticism. The relationship with Religious Mysticism was actually negative. The absorption and magical thinking relationships were all expected and consistent with the scales’ content overlap. The Mysticism pattern was somewhat surprising because I did expect that people who use the PM strategy would also experience mystical feelings in non-religious contexts. However, it seems that this is not the case and that PM spiritual experiences do not necessarily include the existential feelings covered by the Mysticism scale. The negative association of Religious Mysticism with PM is consistent with the weak (sometimes negative) correlation between PM and religiosity. Religious mysticism has been shown to be positively correlated with increased religious satisfaction (Byrd, Lear, & Schwenka, 2000) and it’s unlikely that someone who regularly has these types of experiences would reject their religious belief.
Chapter 6: General Discussion and Conclusion

The goals of this dissertation were to introduce a theoretical construct called ‘need for mystery’ and create the nMyst scale to measure that construct. After elaborating my ideas in the introduction, I followed with a series of four studies describing the item development, structural analyses, and individual item functioning. Reliability was evaluated with both internal consistency and test-retest methods. I linked nMyst to a number of key constructs and criterion variables to establish convergent validity. To establish discriminant validity, I included and ruled out several potentially redundancies. In the following section, I provide an overall summary of the results as they relate to these goals.

Scale Structure. I began Study 1 with a set of 32 items based on the initial six facets that define the nMyst construct. The empirical structure of these items was examined with an exploratory factor analysis. The parallel analysis method suggested an oblique (correlated) three factor solution. Factor 1 included items from facets that focused on rejection of concrete explanations or reliance on explanations from sources other than scientific knowledge. Accordingly, I labeled this factor “Anti-Science”. Factor 2 included items from facets related to Animism and acceptance of strange spiritual experiences. Factor 3 contained a few items from the Romanticism facet but focused on everyday spiritual experiences. Not surprisingly, the two resulting subscales were highly correlated and a combination composite showed a higher reliability than either one alone. Therefore, I opted to proceed with the combination, labeled “Pro-Magic”.

In Study 2, I analyzed a reduced set of 23 items based on the Study 1 results. The massive size of this dataset was exploited to run a series of confirmatory models with independent subsamples. I began with the initial three factor model from Study 2 and released parameters based on LM Tests until I found an acceptable fit. The released parameters were all
theoretically plausible and the model fit was not significantly worse when tested in another independent subsample. The similar fit of the two and three factor models indicated that the more parsimonious two factor model could also be adjusted to fit the data. The correlation between the 2nd and 3rd factors was also very high. A comparison of European heritage and Chinese heritage samples showed that the three factor model failed to fit in the Chinese sample. Their data were actually best fit with the two factor model (again, combining Factors 2 and 3). Hence the criterion of cross-cultural generalizability also favored the two-factor model.

Together the Study 1 and Study 2 results supported my choice of the two factor structure to create the nMyst subscales. These subscales were more consistent with the theoretical interpretations of the items and exhibited good reliability and validity in later studies. It is also important to keep in mind that I am using the item mean composites as the scores for the analysis and not using factor scores. This is a practical consideration; other people using my scale in their research are unlikely to go to the effort of calculating factor scores and I want all research using the scale to be comparable. Also the factor scores would remove variance among items with correlated errors that I would still consider to be a valuable part of the construct. By using the composites I am including all the variance in the items, much of which may be considered noise in a CFA, but is not inconsistent with the overall construct.

**Item Response Theory (IRT).** The application of IRT in Study 2 was useful in revealing several measurement issues that were not evident in the more traditional Classical Test Theory analyses. For example, I found that the AS items were actually measuring most efficiently one and half standard deviations below the average trait level. This asymmetry might be traced to the fact that the items were generated with my brainstorming collaborators – all of us clearly low on Anti-Science. An item that we found very difficult to endorse might be a widely held belief in the general population. To make them easier to endorse, these items should be
subjected to minor wording changes to address these issues in the future. PM, however had none of these issues and had the lowest measurement error at the midpoint of the latent trait.

The differential item functioning (DIF) results showed that the items functioned slightly better in the Chinese heritage sample. The item discrimination parameters were generally larger, leading to greater peak scale information, and the AS subscale was less skewed to the low scores. These findings provide further evidence that the nMyst scale is valid in both groups and again operates slightly better in the Chinese heritage sample.

**Reliability.** Across all four studies, the nMyst subscales showed solid alpha reliabilities: AS was fairly consistent around $\alpha = .85$ and PM, around $\alpha = .78$. These reliabilities were also quite similar across a wide range of different ethnic groups.

Study 4 evaluated the temporal stability of the subscales using test-retest reliability over a three month period. The values were comparable to those of the internal consistency reliabilities. The sample included both European and Asian Canadians, further supporting the cross-cultural consistency. Its temporal stability provides further evidence for the stability of nMyst as a personality trait. This analysis was based on the 20 item set suggested by the analyses in Study 2, found in Appendix C. Based on the studies in this thesis, I consider this the final item set to be used in future research.

**Validity.** The basis for establishing the nomological net is the set of criterion variables hypothesized to relate to the construct of interest. The focal instrument, nMyst, was designed to measure the tendency toward mysterious views of religious and spiritual beliefs. Because such beliefs are the foundation of people’s worldviews, I expected them to relate to a wide variety of external variables. In the four studies reported here, I evaluated links with a total of 10 self-report personality and attitude scales. Of special importance were links to spiritual beliefs. In
addition to conceptually relevant scales, I related nMyst to several demographic variables to situate it in terms of observable life events.

**Personality and worldview correlates.** I argued that the most important correlates of nMyst would be the beliefs where it plays the most significant role. In every study, the AS subscale was significantly related to general religiosity. In Study 1, high PM scores were related to a broad range of paranormal beliefs, supporting the trait premise of the nMyst theory. Mysterious explanations help maintain beliefs in both religious and paranormal phenomena.

nMyst was also expected to relate to fundamental enduring personality traits, that is, the ‘Big Five’: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience (Costa & McCrae, 1987). Current consensus indicates that these five broad traits capture most of the variance in personality. The relationships among the Big Five and nMyst were replicated in both a student and community sample.

Anti-Science was consistently correlated positively with Agreeableness. I believe this is due to a generally tendency for conformity in anti-science attitudes. This is also consistent with the correlation with political conservatism, which is also related to conforming to the status quo or authority figures. Pro-Magic was consistently correlated with Neuroticism and Openness to Experience. Neuroticism entails an inability to control one’s emotions. These strong emotional reactions could be the source of the strong experiences of Awe and Absorption that high PM people interpret as spiritual experiences. The strong link with Openness to Experience was clearly expected, given that the PM strategy of welcoming new experiences and accepting them as part of one’s worldview.

I also investigated clusters of worldview variables that were primarily expected to correlate with Anti-Science attitudes. Religiosity was included in these analyses because it is
correlated with all of these measures. AS correlated positively with authoritarianism and the moral foundation of Purity.

PM was only negatively correlated with two variables in this set, the Respect for Authority moral foundation and PNS. Both of these findings make sense. Many PM beliefs go against the “normative” beliefs promoted by authority figures. Those high in need for structure want to choose the simplest explanation provided and fit everything into a coherent framework. Many PM explanations are highly complex and overall do not adhere to an overarching worldview structure.

The Pro-Magic variables were primarily related to spiritual experiences: tendency to experience Awe, mysticism, absorption, magical thinking, and paranormal beliefs. Note that the overall theme emerging from this constellation of variables is very close to the general description of someone who uses the pro-magic strategy. They can easily become absorbed into an experience and have feelings of awe and oneness with universe and disconnection from time and space. High scorers on nMyst believe that these experiences are proof of magic and mystery in the universe and interpret experiences that other people would find uninteresting to be important and meaningful. The commonality behind these variables motivated me to capture the underlying construct.

I expected belief in Free Will and Determinism to have a similar relationship with both AS and PM. After all, Free Will and Fatalistic Determinism are both examples of mysterious beliefs: They are unprovable and based on unobservable phenomena such as a soul or outside unknowable agents such as God or fate. They are clearly not mechanistic and rely on teleology, (e.g. “We have souls because it is necessary for free will”). By contrast, belief in Scientific Determinism is based on the systematic collection of information about testable causes such as genetic or social structures: These epitomize the dry, concrete explanations that those high in
nMyst reject. Although both subscales were positively correlated with Fatalistic Determinism, Free Will belief was positively correlated with AS and negatively correlated with PM. The explanation may lie in the fact that AS is generally associated with more traditional religious beliefs that have reconciled the coexistence of God’s omnipotence and free will for centuries. Without the influence of religion, PM is related only to the more magical idea of Fate which should logically contradict belief in Free Will.

**Demographic Variables.** In addition to worldview and personality measures, I examined the relationship of nMyst with several demographic variables. These included religious affiliation, conversion, gender, age, ethnicity, and education.

**Religious Affiliation.** The main demographic of interest was religious affiliation. nMyst beliefs should be related to the particular content of a person’s religious belief, not just whether they are or are not religious. I found several patterns in nMyst differences between religions that are consistent with the basic beliefs of the religion.

The results on religious conversion supported my theory that nMyst incompatibility could lead to changes in religious affiliation. These results were consistent with the weaker hypothesis that, were it true that nMyst could influence conversion, then converts should have nMyst profiles more consistent with those in their chosen religion than their original religion. This effect faltered only for cases in which there may be a different perception of the religion between converts and lifelong believers (e.g. Buddhism) or in situations where the converts were even more extreme in their strategy use (e.g. convert Fundamentalists being even more AS).

A quasi-experiment comparing converts to non-converts after the conversion event does not rule out the possibility that the change in religion caused the change in nMyst. However, there are other pieces of evidence that make the reverse causal theory less plausible. Research on
religious conversion has found little change to trait level personality after conversion with most changes found at the level of characteristic adaptations and life narratives (Rambo, 1993).

In the same data, I found that education had a minimal effect on nMyst levels until the graduate level (a small group). Also those with higher levels of education were not more likely than any other education level to convert to Atheism or Agnosticism. All of these results together provide support for the conceptualization of nMyst as a stable personality trait.

**Other demographics.** I also considered several other demographics considered important for establishing the norms of a scale. No gender differences emerged on either subscale in the student sample. In both community samples, however, women were slightly higher on both subscales. That result is consistent with evidence that women are higher than men in both paranormal and religious belief (Emmons & Sobal, 1981; Pew Research Center, 2007).

PM was found to decline with age, consistent with the correlation between age and paranormal beliefs (Tobacyk, Pritchett, & Mitchell, 1988). AS, however, did not change with age. There is usually a positive relationship between age and religiosity (Pew Research Center, 2007). However, in this sample, age was also uncorrelated with religiosity so this finding may not generalize to other samples.

I had no particular hypotheses regarding mean differences between ethnicities. Instead, my main interest in looking at ethnic differences was to find evidence for cross-cultural generalizability of the nMyst scale. The factors were distinct and reliabilities were comparable across groups, thereby providing evidence for cross-cultural validity.

**Summary**

The research presented in this thesis addresses an array of hypotheses relating to nMyst and its validation as a useful individual difference construct. Overall, the data have generally
supported my interpretation of the AS and PM subscales with most of my hypotheses being confirmed.

The nomological net of nMyst appears coherent in that nMyst is related to the types of beliefs expected, religious and paranormal, and also related to generally preferring teleological explanations like Free Will or Fate. AS is related to social conservative views and morality, but not necessarily related to a particular cognitive style, such as PNS. PM is related to tendencies to have spiritual experiences or affective reactions that are interpreted as spiritually meaningful. For the most part, these findings support my theory of nMyst and interpretations of the subscales. With a bit of adjustment to the AS subscale items; nMyst could be a useful measure to address various individual differences in religious beliefs and spirituality within and across cultures.

The nature of Need for Mystery: Need, Trait, or Coping Style? Several pieces of evidence support the classification of nMyst as both a need and a personality trait. Although the desire to maintain and protect cherished beliefs is a universal human tendency, only high nMyst scorers reach the level of a need. The mystification of their beliefs makes them unassailable. Perhaps they are especially sensitive about preserving their beliefs: Therefore they put in more effort to protect their beliefs and satisfy their need. Alternatively, they have discovered the coping power of mystification and opted to protect their beliefs that way.

The lengths that high nMyst scorers will go to satisfy this need is evidenced by their willingness to convert religions. Longitudinal evidence of nMyst preceding conversion is still necessary, but thus far the theory has been supported. Another piece of evidence to support the need status of nMyst is that its satisfaction results in positive affect (see Sheldon, Elliot, Kim, & Kasser, 2001). This question could be answered with a study showing a positive affective reaction to reading accounts of paranormal encounters (for Pro-Magic), or writing a rebuttal to a scientific threat to a cherished belief (for Anti-Science).
The cross-cultural reliability of the nMyst scale provides further support for nMyst as a fundamental psychological need. One proposed criterion for a fundamental psychological need is that it is universal (Baumeister & Leary, 1995). Religious and supernatural beliefs are human universals, based on evolved tendencies in the human mind (Atran & Norenzayan, 2004). If nMyst represents individual differences in a need for these beliefs, then it should also be universal following the beliefs is needs and protects.

My data also support the notion that nMyst is a stable personality trait. Its temporal stability over three months is comparable to its internal consistency reliability. Very little additional error is introduced by having people respond at different times. If nMyst were more state-like or under conscious control one would expect additional variance across measurement occasions.

The second component of a personality trait is consistency across domains. Obviously, given the high correlation between the subscales, people high in nMyst are usually able to apply both strategies to protect their beliefs. High-high people may use AS to protect their paranormal beliefs as well as religious beliefs. They also supplement and protect their religious belief by incorporating other beliefs. I would also expect that someone high in nMyst would apply protection techniques to any of their deeply held beliefs even outside the domains of the paranormal or religious. Pseudoscience beliefs are not intended to be paranormal, but have no basis in observable science and can easily be protected using nMyst. Showing nMyst strategy application will be a primary goal in future research. Showing which beliefs provide either a protection against threat or provoke a rejection reaction will give insight into just how broadly people can apply nMyst.

Another possible aspect of nMyst is its function as a way of coping with uncertainty in the world. Worldview threats cause anxiety and by making their worldview “mysterious”, people
can protect themselves from this anxiety. Beliefs that are not based on evidence are more stable and safe from disconfirming evidence. The two strategies, AS and PM, are specific manifestations of how to go about making one’s worldview mysterious. AS is more reactionary and is used in response to a direct perceived threat. PM is more protective and acts as a buffer against possible future threats.

**Limitations.** Unsurprisingly, there are limitations in these studies that should be addressed in further research. My hope that the proposed item set would fall neatly into a simple factor structure was not met. Instead, the evidence waffled between two and three factors. There was also some evidence that the factor structure differed across cultural groups. Future exploration of the factor structure could include item parceling. This approach often yields more clear and replicable factor structures. Nonetheless, my ultimate goal, that is, to create a useful pair of nMyst subscales appears to have been met. They fit reasonably into a nomological network of constructs and seem to produce predictable results. However, the interplay between theorizing and item creation likely requires some additional refinements to more directly address some theoretical components, such as teleology.

On the whole, the items themselves performed well -- according to the size of factor loadings and IRT analyses. One issue of concern is the low difficulty of the AS subscale items. This issue could be addressed in future studies by adjusting the wording of the items to make them more extreme in tone. However, it is always possible that this adjustment could lead to changes in the overall scale structure. Any changes would necessitate a new CFA analysis and IRT to confirm that the desired change in difficulty was achieved without damaging the structure.

Another important limitation is the self-report and cross-sectional nature of the studies. Simple self-report correlational studies are a necessary foundation for establishing a new
construct. However, informant ratings and behavioral correlates are essential to demonstrating the effects of nMyst extend beyond the intra-psychic level (Cohrs, Kämpfe-Hargrave, & Riemann, in press). An additional complication to these particular self-report data is that they were all collected online. Online studies have been criticized since psychologist first had the idea to use the technology and several papers have responded to criticisms that online participants are unmotivated or somehow different from in lab participants (Gosling, et al., 2004). In all studies included in this thesis, I included “catch questions” to weed out random respondents and paid an amount on MTurk that seems to be ideal for encouraging interested participants (Burhmester, Kwang, & Gosling, 2011). Online data collection has potential downsides; however, psychologists are increasingly discovering that having a larger and more representative sample outweighs any issues and most concerns have generally been unfounded.

I have a number of ideas for future studies using nMyst now that I have laid the basic foundation for the theory and measure. I explore these in the following section.

**Future Directions and Predictions**

There are several areas of psychology where nMyst could make a significant research contribution. Below, I give an overview of three such domains – the Meaning Maintenance Model, meaning in life, and suggestibility – and speculate on how nMyst could contribute.

**Meaning Maintenance Model.** According to the Meaning Maintenance Model, people create meaning frameworks from associations among concepts and use these frameworks to make sense of the world (Heine, Proulx, & Vohs, 2006). When an important assumption is violated, a meaning threat is triggered. People are motivated by such threat responses to reaffirm their sense of meaning and several types of reaffirmation have been tested.

The types of meaning violations are widely varied. The first account of the meaning threat phenomenon was found in response to exposure to playing cards in which the colors of the
suits had been switched (i.e. black hearts and red spades; Bruner and Postman, 1949). Participants did not notice the switch at first, but gradual came to feel that something was wrong. In some cases they were never able to consciously realize what was wrong with the cards and those people experienced distress at being unable to pinpoint the source of their discomfort. Much later, a similar effect was discovered as a response to people being reminded of their own death. This line of research became known as Terror Management Theory, which considered a fear of death to be the primary source of threat and not a violation of meaning (Pyszczynski, Greenberg, & Solomon, 2004; Solomon, Greenberg, & Pyszczynski, 1991).

The Meaning Maintenance Model is broader than TMT, including both meaning violations and mortality salience. It includes violations of associations in the external world (e.g. the playing cards or absurdist literature; Bruner & Postman, 1949; Proulx & Heine, 2009), associations within the self (exemplified by Cognitive Dissonance; Festinger, 1957), and associations between the self and the external world (such as one’s cultural identity; Solomon, Greenberg, & Pyszczynski, 1991).

I hypothesize that the mysterious beliefs of those high in nMyst play an important role in meaning maintenance. By making beliefs unassailable, they help maintain meaning and order for confusing or seemingly random occurrences. nMyst beliefs could be seen as including associations within the external world and associations between the self and the external world. As described in the introduction, mysterious beliefs are generally existential, imparting important facts about the world that can also relate to the individual. Threats or violations to these beliefs should cause a meaning threat reaction under the MMM framework.

In response to a meaning threat, people generally respond by affirming an association in another domain using fluid compensation (Heine, Proulx, & Vohs, 2006). Fluid compensation is a theory that states that there is an overall sense of meaning made up of all the meaningful
associations included in a person’s worldview. If one of these sources of meaning is threatened, they can repair their overall sense of meaning by bolstering or affirming another association. The original measure of the effect was increased bail set for prostitutes after judges were exposed to a morality salience manipulation (Rosenblatt, Greenberg, Solomon, Pyszczynski, & Lyon, 1989). The increase in punishment for a crime that is socially defined has been used as a measure of meaning compensation in many TMT studies and meaning maintenance studies with all sorts of meaning violations (e.g. Proulx, Heine, & Vohs, 2010). Other measures of cultural adherence include increased aggression toward those who criticize the in-group (Pyszczynski et al., 2006) and increase affirmation of one’s own cultural identity (Proulx, Heine, & Vohs, 2010).

Another way of reaffirming one’s sense of meaning is creating meaning by making a new associations (Proulx & Heine, 2009; Randles, Proulx, & Heine, 2011). Instead of affirming another aspect of meaning, such as cultural associations, people temporarily become better at finding new associations. This effect was measured using an implicit learning task. The advantage of the implicit learning measure of meaning threat is that it is independent of prior attitudes. Punishing a prostitute is only an affirmation of meaning if one ascribes to the view that prostitutes are violating a meaningful part of one’s culture. This view is going to be confounded with social conservative attitudes, making it more difficult to use in individual differences contexts in which conservatism is related to the variables of interest, such as nMyst.

The basic theory that nMyst beliefs are an important part of a person’s worldview can be tested using a meaning maintenance paradigm. Meaning can be violated by presenting evidence against a particular mysterious belief and the threat response can be measured using an implicit learning task. Meaning maintenance would be an ideal paradigm for testing domains in which nMyst predicts a meaning threat reaction. Demonstrating that those high in nMyst experience
threat in response to a wide variety of beliefs would provide further evidence for nMyst as a personality trait.

It may also be possible to distinguish between the two factors using the MMM. It is possible that people who are high AS, but low on PM would not have their threat resolved by the PM strategy and vice versa. People who are high in both factors should have their threat resolved in either case. People low in nMyst should not show a threat response when a mysterious belief is undermined. Specifically, PM should be protective and perhaps those high in PM would not even experience a threat (or experience a weaker threat reaction) because their strategy is in place before the threat occurs. AS on the other hand is reactive and is used in response to a threat. So high AS only people should experience a strong threat and have that threat resolved by denouncing the source or basis for the threat.

**Meaning in life.** The emerging field of positive psychology focuses on improving people’s lives through finding out what makes them happy and increases wellbeing. One of the general findings of this research has been that having a sense of meaning in one’s life is associated with greater wellbeing (Mascaro & Rosen, 2005; Reker, Peacock, & Wong, 1987; Ryff, 1989; Steger & Frazier, 2005; Zika & Chamberlain, 1987, 1992). Another line of research has found that increasing positive mood, through manipulation, also increases a sense of meaning in life (King, Hicks, Krull, & Del Gaiso, 2006). However, the effect is mediated by several other variables that can be sources of meaning, specifically religiosity and Faith in Intuition (Hicks & King, 2007; Hicks et al., 2010). People who were higher in these individual difference variables showed both a higher baseline score on a meaning in life measure and did not show an increase in meaning in response to a mood manipulation. King and colleagues conclude that mood can be an indicator of meaning, but only in the absence of a more stable source of meaning, such as religious belief.
I believe that nMyst would work in the same way to provide a stable sense of meaning using either strategy. AS is almost certain to mediate the effect of mood on meaning given its association with religiosity. I am more interested in testing that PM has the same effect because that would provide evidence that supernatural and paranormal beliefs provide the same meaning benefits as more traditional religious belief. People who do not share these beliefs often hold them in contempt, which is especially clear in some of the previous research on paranormal beliefs (e.g. Sheep-Goat scales in which believers are deemed “Sheep”, Thalbourne, 1981; impaired reasoning ability, Weirzbicki, 1985). Though some outspoken atheists and skeptics may hold similar feelings about religious belief, the importance of being respectful of those beliefs in academic research is strongly stressed. Some paranormal belief researchers have suggested that these beliefs provide a way of coping with traumatic life events (Irwin, 1994; 2000). However, this relationship with coping style was not found in subsequent empirical tests (Rogers, Qualter, Phelps, & Gardner, 2006). Perhaps, rather than filling a similar coping role to religious belief, paranormal belief may fill a similar meaning creating role. If paranormal beliefs provide the same kind of meaning that religion provides and are equally meaningful to those who hold those beliefs, it could help to legitimize alternative beliefs that are currently marginalized.

**Suggestibility.** Psychologists have long been interested in explain why people hold paranormal beliefs. A series of experiments established that believers tend to engage in selective learning (Greenwald & Sakamura, 1967) and only attend to evidence that confirms their beliefs (Jones & Russell, 1980). Believers’ ratings of their own ESP were unrelated to their feedback on a card guessing task; in contrast, skeptics’ ratings were highly related to the feedback. Those who hold paranormal beliefs also show consistent logic errors when evaluating the truth of a hypothesis based on evidence (Wierzbicki, 1985). They were more likely to make the error of affirming the consequent. When given a hypothesis framed as “If A then B”, and told that B is
true, they conclude that A must be true. When testing the hypothesis that “if telekinesis is possible, then I can move that pen without touching it”, movement of the pen is considered to be proof of telekinesis without considering alternative explanations for the movement.

Though logical fallacies may be related to maintaining paranormal beliefs, Pro-Magic should predict a general susceptibility to adopting new beliefs. Magical thinking is the belief that a person can use their consciousness to alter reality, such as praying for someone to recover from illness (Nemeroff & Rozin, 2000; Subbotsky, 2007). The magical nature of such rituals is usually consciously recognized as illogical, but is appealing on an emotional level. The duality of this effect has been termed “participation” (Lévy-Bruhl, 1966). People can participate in a ritual and act as if the effect is true on some level without explicitly believing it. In one study, participants labeled an empty bottle either as “sucrose” or “sodium cyanide, poison”. The bottles were then filled with sugar water and participants were asked to drink out of the bottles. Participants preferred to drink from the bottle labeled “sucrose” even though they knew that the bottle labeled “poison” had the same contents (Nemeroff & Rozin, 2000). Their participation in labeling the bottle as poison affected their emotional response to drinking out of the bottle despite being consciously aware that it was not actually poison.

Participation effects can make people more susceptible to manipulation and suggestion even when the suggestion does not make rational sense (Subbotsky, 2007). For those high in nMyst, who greatly value beliefs that cannot be shaken by “proof”, these types of suggestions may be even more powerful. Especially if a particular suggestion is very intuitively appealing. These suggestions can lead to fairly harmless beliefs, such as astrology, or very harmful beliefs like homeopathic medicine.

Homeopathic medicine is a type of “alternative medicine” invented in the 18th century. The basic principal is that if a substance can cause a particular symptom then a very dilute
solution of that substance can be used to cure a disease with the same symptom. The dilution is often so great that there is no molecular trace of the original “cure” in the final preparation; however, the theory claims that a more dilute solution is actually more potent than a less dilute solution. All reviews of clinical trials have found no effect of homeopathic medicine greater than a placebo effect (Ernst, 2002). Despite the lack of evidence of effectiveness, several European countries cover homeopathy in their health care systems and homeopathic remedies are sold all over the world (WHO, 2001).

It seems likely that participation and magical thinking are at the root of this phenomenon. Only the skeptics are even concerned with the efficacy of the remedies. James Randy, the professional skeptic, has been very vocal in trying to debunk homeopathy (Jones, 2010). Homeopathic doctors and patients who have decided to participate consciously choose, on an emotional level, to take the remedies and believe that they work. The emotional attachment to the theory interacts with distrust in science to bolster the perception that the remedy is working.

Distrust in science is another large part of susceptibility to paranormal beliefs or alternative medicines. Ridolfo, Baxter, & Lucas (2010) tested the interaction between popularity and scientific support in paranormal beliefs. They correctly hypothesized that paranormal beliefs that were presented as popular with the general public would be believed more than those presented as unpopular. They also expected that if a belief was unpopular, then scientific support would lead to more belief. What they actually found was, regardless of the popularity, scientific support actually lead to less belief. Participants were more likely to believe an unpopular paranormal belief that wasn’t supported by science than an unpopular belief that was supported. This result suggests a general distrust of science; however, this study did not account for individual differences. I believe that people who are low in nMyst, may not be swayed by the popularity of the idea, but may report an increase in belief for paranormal ideas supported by
science. It would make sense for an aggregate study to find the high nMyst pattern because there are fewer individuals in the low-low quadrant than in the other three.

Though Anti-Science is one part of nMyst, there are people who are high in Pro-Magic who do not immediately dismiss scientific evidence. These people may still show the same suggestibility as people high in both factors, but may also be more easily convinced of scientific findings and less able to accept disproof than those low in nMyst.

One excellent example is the reported relationship between vaccination and autism. Clearly, in order to believe the finding in first place, people had to be accepting of medical science in general (in fact, mothers of children who are not vaccinated are more likely to have a college education; Smith, Chu, & Barker, 2004). Unfortunately, they were too quick to accept the evidence and then incorporated it into their worldview. Believing that vaccinations are bad may be intuitively easy to accept because they cause physical pain and many people do not really understand how they work. Parents’ refusals of vaccinations lead to outbreaks of diseases in several states, including measles and pertussis (Omer et al., 2009). It has since been discovered that the link between vaccines and autism was actually not true (Plotkin, Gerber, & Offit, 2009), but many parents invested in the idea as an easy solution and are now refusing to believe that they were lied to.

Excessive suggestibility can prove dangerous in a variety of contexts. It is important to understand why some people are more gullible than others to try to prevent and combat the spread of misinformation. I have found that those high in PM find common paranormal beliefs and conspiracy theories to be more plausible. A stronger test of general suggestibility would be to fabricate a paranormal belief or conspiracy theory and have participants rate how plausible they find the suggested beliefs. Similar to the Ridolfo, Baxter, & Lucas (2010) study, I would also vary the popularity of the phenomenon and the scientific backing.
In attempting interventions in anti-science thinking and suggestibility it is important to understand nMyst as a psychological need. Treating these beliefs with contempt and not acknowledging what they provide for believers likely causes a greater threat response. Any attempt to change beliefs like homeopathy need to be undertaken with respect for the believers. Perhaps, providing mysterious but not harmful alternatives would be more successful. Scientific findings could also be presented in a more appealing and mysterious way. If evolution is presented in a way that emphasizes how unlikely and amazing life is, implying that it must have some great purpose, people were more likely to endorse it after a meaning threat (Tracy, Hart, & Martens, 2011).

**Final Comments.** I have argued that the concept of need for mystery can help explain why some people prefer fuzzy, unsubstantiated beliefs over more concrete scientific ones. Although application of the term “need” may be problematic for some readers, I have argued that it is justified in the case of mysterious beliefs. In brief, the four studies detailed in this dissertation indicate that individual differences in nMyst can be measured.

Note that I am not recommending interventions to challenge the belief structures of those high in nMyst. As a coping mechanism, nMyst provides psychological benefits to those individuals. Respect for the idiosyncratic meaning constructs of others is important for both research and possible interventions. Both components may promote happiness and well-being. Nonetheless, I believe that society as a whole would benefit from a decrease in anti-science attitudes. Except in unusual circumstances, pro-magic beliefs may be no more harmful than other forms of intellectual exploration.

My speculations on future nMyst research included further validation as well as a breadth of potential applications to grander issues of worldview beliefs and personal meaning. The latter issues form a large part of both social and personality psychology. I have argued that paying
attention to individual differences in nMyst can help clarify the dramatic variation seen in the way such beliefs are held.

I also argued that the nMyst scale will be an especially fruitful instrument for cross-cultural survey research: It captures two facets of a culturally-important individual difference variable with only a small number of items. Of course, the manifestations of nMyst are sure to differ in interesting ways across cultures. In conclusion, I believe that nMyst captures a useful and provocative personality variable and the nMyst scale has great potential for the future study of beliefs.
### Tables

Table 1. Factor loadings of the initial 32 item set in Study 1.

<table>
<thead>
<tr>
<th>Item</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>.122</td>
<td>.337</td>
<td>-.026</td>
</tr>
<tr>
<td>Item 2</td>
<td>.021</td>
<td>-.025</td>
<td>-.054</td>
</tr>
<tr>
<td>Item 3</td>
<td>.528*</td>
<td>-.012</td>
<td>-.394</td>
</tr>
<tr>
<td>Item 4</td>
<td>-.125</td>
<td>.338</td>
<td>-.233</td>
</tr>
<tr>
<td>Item 5</td>
<td>.317</td>
<td>.128</td>
<td>.072</td>
</tr>
<tr>
<td>Item 6</td>
<td>.853*</td>
<td>-.088</td>
<td>-.116</td>
</tr>
<tr>
<td>Item 7</td>
<td>-.312</td>
<td>-.021</td>
<td>-.080</td>
</tr>
<tr>
<td>Item 8</td>
<td>.672*</td>
<td>-.135</td>
<td>-.019</td>
</tr>
<tr>
<td>Item 9</td>
<td>-.399</td>
<td>-.180</td>
<td>.050</td>
</tr>
<tr>
<td>Item 10</td>
<td>.800*</td>
<td>.151</td>
<td>-.152</td>
</tr>
<tr>
<td>Item 11</td>
<td>.718*</td>
<td>-.036</td>
<td>-.174</td>
</tr>
<tr>
<td>Item 12</td>
<td>-.502*</td>
<td>-.113</td>
<td>-.031</td>
</tr>
<tr>
<td>Item 13</td>
<td>.709*</td>
<td>-.149</td>
<td>.065</td>
</tr>
<tr>
<td>Item 14</td>
<td>.177</td>
<td>.034</td>
<td>.044</td>
</tr>
<tr>
<td>Item 15</td>
<td>-.012</td>
<td>.587*</td>
<td>-.061</td>
</tr>
<tr>
<td>Item 16</td>
<td>.011</td>
<td>-.456*</td>
<td>.058</td>
</tr>
<tr>
<td>Item 17</td>
<td>-.091</td>
<td>-.026</td>
<td>.300</td>
</tr>
<tr>
<td>Item 18</td>
<td>-.007</td>
<td>.724*</td>
<td>-.039</td>
</tr>
<tr>
<td>Item 19</td>
<td>.065</td>
<td>.146</td>
<td>.236</td>
</tr>
<tr>
<td>Item 20</td>
<td>.193</td>
<td>.478*</td>
<td>.208</td>
</tr>
<tr>
<td>Item 21</td>
<td>-.619*</td>
<td>.158</td>
<td>-.082</td>
</tr>
<tr>
<td>Item 22</td>
<td>.502*</td>
<td>.214</td>
<td>.105</td>
</tr>
<tr>
<td>Item 23</td>
<td>.277</td>
<td>.061</td>
<td>.246</td>
</tr>
<tr>
<td>Item 24</td>
<td>.326</td>
<td>.034</td>
<td>.332</td>
</tr>
<tr>
<td>Item 25</td>
<td>.069</td>
<td>-.196</td>
<td>.447*</td>
</tr>
<tr>
<td>Item 26</td>
<td>-.048</td>
<td>.008</td>
<td>.532*</td>
</tr>
<tr>
<td>Item 27</td>
<td>.180</td>
<td>.411*</td>
<td>.265</td>
</tr>
<tr>
<td>Item 28</td>
<td>.252</td>
<td>.076</td>
<td>.632*</td>
</tr>
<tr>
<td>Item 29</td>
<td>.053</td>
<td>-.024</td>
<td>.314</td>
</tr>
<tr>
<td>Item 30</td>
<td>-.016</td>
<td>.266</td>
<td>.269</td>
</tr>
<tr>
<td>Item 31</td>
<td>.159</td>
<td>.441*</td>
<td>.167</td>
</tr>
<tr>
<td>Item 32</td>
<td>-.020</td>
<td>.253</td>
<td>.184</td>
</tr>
</tbody>
</table>

*Note. N = 217. Pattern matrix results, ML extraction with oblimin rotation. * Item loads strongly on this factor.*
Table 2. Final item set loadings, Study 1.

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1. I prefer not to know how magic tricks are done.</td>
<td>.110</td>
<td>.329</td>
<td>-.047</td>
</tr>
<tr>
<td>Item 4. Love is just a chemical reaction.</td>
<td>-.115</td>
<td>.274</td>
<td>-.143</td>
</tr>
<tr>
<td>Item 5. There are some things that humans should not try to understand.</td>
<td>.297*</td>
<td>.104</td>
<td>.080</td>
</tr>
<tr>
<td>Item 7. I find it hard to believe in things that I can't picture.</td>
<td>-.277</td>
<td>-.061</td>
<td>-.078</td>
</tr>
<tr>
<td>Item 8. There are things in this world too complicated to happen on their own.</td>
<td>.666*</td>
<td>-.140</td>
<td>-.031</td>
</tr>
<tr>
<td>Item 9. Reports of so-called 'death experiences' don't prove there is an afterlife.</td>
<td>-.448*</td>
<td>-.147</td>
<td>.074</td>
</tr>
<tr>
<td>Item 10. I prefer spiritual explanations for things rather than scientific ones.</td>
<td>.729*</td>
<td>.118</td>
<td>-.060</td>
</tr>
<tr>
<td>Item 12. If someone told me they had a spiritual vision I would doubt it.</td>
<td>-.593*</td>
<td>-.052</td>
<td>.064</td>
</tr>
<tr>
<td>Item 13. I believe there is something beyond this material universe.</td>
<td>.762*</td>
<td>-.179</td>
<td>.000</td>
</tr>
<tr>
<td>Item 15. I sometimes think that inanimate things are fighting against me (e.g., computers, cars).</td>
<td>-.053</td>
<td>.591*</td>
<td>-.016</td>
</tr>
<tr>
<td>Item 16. I never think of objects as having a personality.</td>
<td>-.033</td>
<td>-.467*</td>
<td>.078</td>
</tr>
<tr>
<td>Item 17. Animals have souls too.</td>
<td>.004</td>
<td>-.026</td>
<td>.123</td>
</tr>
<tr>
<td>Item 18. Objects sometimes move around on their own.</td>
<td>-.046</td>
<td>.749*</td>
<td>.011</td>
</tr>
<tr>
<td>Item 20. Some houses are definitely haunted with ghosts.</td>
<td>.216</td>
<td>.443*</td>
<td>.207</td>
</tr>
<tr>
<td>Item 21. I am comfortable with the idea that the world is just atoms and molecules.</td>
<td>-.672*</td>
<td>.192</td>
<td>-.011</td>
</tr>
<tr>
<td>Item 22. I trust my intuition more than scientific evidence.</td>
<td>.508*</td>
<td>.188</td>
<td>.060</td>
</tr>
<tr>
<td>Item 24. I sometimes get a shiver thinking about mysterious things.</td>
<td>.248</td>
<td>.071</td>
<td>.343*</td>
</tr>
<tr>
<td>Item 25. Looking up at the stars thrills me.</td>
<td>-.023</td>
<td>-.161</td>
<td>.525*</td>
</tr>
<tr>
<td>Item 26. I sometimes get a shiver hearing certain musical pieces.</td>
<td>-.158</td>
<td>.080</td>
<td>.653*</td>
</tr>
<tr>
<td>Item 27. When I'm on my own, I sometimes get magical feelings.</td>
<td>.039</td>
<td>.503*</td>
<td>.382</td>
</tr>
<tr>
<td>Item 28. I sometimes get goosebumps when I'm really connecting with people.</td>
<td>.172</td>
<td>.147</td>
<td>.611*</td>
</tr>
<tr>
<td>Item 29. I often get the feeling that I've done something before - even though it's not possible (deja vu).</td>
<td>-.010</td>
<td>-.005</td>
<td>.345*</td>
</tr>
<tr>
<td>Item 31. Some people can predict the future.</td>
<td>.169</td>
<td>.401*</td>
<td>.155</td>
</tr>
</tbody>
</table>

Note. N = 217. Pattern matrix results, ML extraction with oblimin rotation. *Item loads on this factor.
Table 3. CFA Model fit comparisons from Study 2.

<table>
<thead>
<tr>
<th>Model Description</th>
<th>$\chi^2$</th>
<th>df</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA [CI .90]</th>
<th>$\chi^2$ Diff(df)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1 (3 Factor, No errors or cross-loadings)</td>
<td>613.58</td>
<td>186</td>
<td>.85</td>
<td>.83</td>
<td>.07 [.06 - .08]</td>
<td></td>
</tr>
<tr>
<td>Model 2 (3 Factor, with cross-loadings)</td>
<td>482.27</td>
<td>181</td>
<td>.90</td>
<td>.85</td>
<td>.06 [.05 - .07]</td>
<td>131.31(5) Model 1</td>
</tr>
<tr>
<td>Model 3 (2 Factor, No errors or cross-loadings)</td>
<td>653.25</td>
<td>188</td>
<td>.84</td>
<td>.82</td>
<td>.07 [.07 - .08]</td>
<td>39.67(2) Model 1</td>
</tr>
<tr>
<td>Model 4 (2 Factor, with cross-loadings)</td>
<td>529.61</td>
<td>184</td>
<td>.88</td>
<td>.86</td>
<td>.06 [.05 - .07]</td>
<td>123.64 (4) Model 3</td>
</tr>
<tr>
<td>Model 5 (3 Factor, with errors and cross-loadings)</td>
<td>361.89</td>
<td>177</td>
<td>.95</td>
<td>.93</td>
<td>.05 [.04 - .05]</td>
<td>120.38(4) Model 2</td>
</tr>
</tbody>
</table>

*Note.* Models were tested using the software EQS (Bentler, 2004). All Satorra-Bentler $\chi^2$ statistics and DIF tests are significant at $p < .001$. 
Table 4. IRT discrimination parameters from the two dimensional model and difficulty parameters from the unidimensional models.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$a1$</th>
<th>$a2$</th>
<th>$b1$</th>
<th>$b2$</th>
<th>$b3$</th>
<th>$b4$</th>
<th>$SS \chi^2 (df)$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 5</td>
<td>1.19</td>
<td>0</td>
<td>-1.47</td>
<td>-0.42</td>
<td>0.01</td>
<td>1.45</td>
<td>173.4 (165)</td>
<td>0.31</td>
</tr>
<tr>
<td>Item 8</td>
<td>2.05</td>
<td>0</td>
<td>-2.83</td>
<td>-2.04</td>
<td>-1.29</td>
<td>-0.3</td>
<td>167.4 (135)</td>
<td>0.03</td>
</tr>
<tr>
<td>Item 9</td>
<td>1.98</td>
<td>0</td>
<td>-1.48</td>
<td>-0.4</td>
<td>0.3</td>
<td>1.32</td>
<td>137.2 (149)</td>
<td>0.75</td>
</tr>
<tr>
<td>Item 10</td>
<td>2.92</td>
<td>0</td>
<td>-1.35</td>
<td>-0.6</td>
<td>0.38</td>
<td>1.1</td>
<td>167.6 (138)</td>
<td>0.04</td>
</tr>
<tr>
<td>Item 12</td>
<td>1.53</td>
<td>0</td>
<td>-2.72</td>
<td>-0.9</td>
<td>0.06</td>
<td>1.78</td>
<td>173.5 (164)</td>
<td>0.29</td>
</tr>
<tr>
<td>Item 13</td>
<td>2.02</td>
<td>0</td>
<td>-4.49</td>
<td>-3.41</td>
<td>-2.32</td>
<td>-0.9</td>
<td>140 (111)</td>
<td>0.03</td>
</tr>
<tr>
<td>Item 15</td>
<td>0</td>
<td>0.92</td>
<td>0.15</td>
<td>1.38</td>
<td>2.12</td>
<td>4.31</td>
<td>157 (132)</td>
<td>0.07</td>
</tr>
<tr>
<td>Item 16</td>
<td>-0.2</td>
<td>0.74</td>
<td>-1.57</td>
<td>-0.16</td>
<td>1.07</td>
<td>3.92</td>
<td>165.4 (168)</td>
<td>0.54</td>
</tr>
<tr>
<td>Item 17</td>
<td>-0.21</td>
<td>1.29</td>
<td>-1.94</td>
<td>-1.16</td>
<td>-0.24</td>
<td>0.96</td>
<td>192.5 (166)</td>
<td>0.08</td>
</tr>
<tr>
<td>Item 18</td>
<td>0</td>
<td>1.62</td>
<td>-0.18</td>
<td>0.6</td>
<td>1.39</td>
<td>2.86</td>
<td>132.6 (120)</td>
<td>0.20</td>
</tr>
<tr>
<td>Item 19</td>
<td>0</td>
<td>0.78</td>
<td>-1.69</td>
<td>-0.24</td>
<td>1.01</td>
<td>3.07</td>
<td>199 (171)</td>
<td>0.07</td>
</tr>
<tr>
<td>Item 20</td>
<td>0</td>
<td>1.99</td>
<td>-1.16</td>
<td>-0.45</td>
<td>0.22</td>
<td>1.19</td>
<td>129.6 (144)</td>
<td>0.80</td>
</tr>
<tr>
<td>Item 21</td>
<td>1.56</td>
<td>0</td>
<td>-2.35</td>
<td>-1.31</td>
<td>-0.49</td>
<td>0.46</td>
<td>204.5 (163)</td>
<td>0.02</td>
</tr>
<tr>
<td>Item 22</td>
<td>1.97</td>
<td>0</td>
<td>-2.07</td>
<td>-0.98</td>
<td>-0.02</td>
<td>1.06</td>
<td>145 (152)</td>
<td>0.64</td>
</tr>
<tr>
<td>Item 23</td>
<td>0.76</td>
<td>0</td>
<td>-5.53</td>
<td>-3.14</td>
<td>-1.63</td>
<td>1.72</td>
<td>150.1 (142)</td>
<td>0.31</td>
</tr>
<tr>
<td>Item 24</td>
<td>0</td>
<td>1.54</td>
<td>-1.76</td>
<td>-0.87</td>
<td>-0.25</td>
<td>1.35</td>
<td>144.1 (155)</td>
<td>0.72</td>
</tr>
<tr>
<td>Item 25</td>
<td>0</td>
<td>0.87</td>
<td>-3.91</td>
<td>-2.19</td>
<td>-0.83</td>
<td>1.3</td>
<td>125.4 (107)</td>
<td>0.11</td>
</tr>
<tr>
<td>Item 27</td>
<td>0.08</td>
<td>1.56</td>
<td>-0.96</td>
<td>-0.09</td>
<td>0.61</td>
<td>1.86</td>
<td>162.9 (144)</td>
<td>0.13</td>
</tr>
<tr>
<td>Item 28</td>
<td>0</td>
<td>1.2</td>
<td>0.23</td>
<td>1.03</td>
<td>1.55</td>
<td>2.86</td>
<td>193.1 (158)</td>
<td>0.03</td>
</tr>
<tr>
<td>Item 29</td>
<td>0</td>
<td>1.04</td>
<td>-3.31</td>
<td>-2.22</td>
<td>-1.16</td>
<td>0.79</td>
<td>158.9 (148)</td>
<td>0.26</td>
</tr>
<tr>
<td>Item 31</td>
<td>0</td>
<td>1.64</td>
<td>-1.61</td>
<td>-0.63</td>
<td>0.13</td>
<td>1.86</td>
<td>129.8 (129)</td>
<td>0.47</td>
</tr>
<tr>
<td>Item 32</td>
<td>0</td>
<td>0.77</td>
<td>-2.93</td>
<td>-1.35</td>
<td>0.17</td>
<td>3.12</td>
<td>182 (150)</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Note. Item numbers correspond to the original item set. Alpha discrimination parameters estimated from a two dimensional IRT model using Bock-Aitkin EM (Bock & Aitkin, 1981) estimation in IRTPRO. Difficulty parameters (b) were estimated in a unidimensional model using the same estimation procedure and software. Item fit parameters are based on the two dimensional model.
Table 5. IRT parameters for the Chinese heritage sample and DIF measures in comparison to the European-American sample.

<table>
<thead>
<tr>
<th>Item</th>
<th>Parameters</th>
<th>DIF Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$a$</td>
<td>$b_1$</td>
</tr>
<tr>
<td>Item 5</td>
<td>1.28</td>
<td>-1.9</td>
</tr>
<tr>
<td>Item 8</td>
<td>1.26</td>
<td>-3.14</td>
</tr>
<tr>
<td>Item 9</td>
<td>1.27</td>
<td>-1.92</td>
</tr>
<tr>
<td>Item 10</td>
<td>2.75</td>
<td>-1.54</td>
</tr>
<tr>
<td>Item 12</td>
<td>1.39</td>
<td>-2.26</td>
</tr>
<tr>
<td>Item 13</td>
<td>1.65</td>
<td>-3.69</td>
</tr>
<tr>
<td>Item 15</td>
<td>1.69</td>
<td>-0.04</td>
</tr>
<tr>
<td>Item 16</td>
<td>0.87</td>
<td>-1.3</td>
</tr>
<tr>
<td>Item 17</td>
<td>1.41</td>
<td>-2.04</td>
</tr>
<tr>
<td>Item 18</td>
<td>1.71</td>
<td>-0.31</td>
</tr>
<tr>
<td>Item 19</td>
<td>0.97</td>
<td>-2.84</td>
</tr>
<tr>
<td>Item 20</td>
<td>1.76</td>
<td>-1.16</td>
</tr>
<tr>
<td>Item 21</td>
<td>1.2</td>
<td>-2.06</td>
</tr>
<tr>
<td>Item 22</td>
<td>2.16</td>
<td>-2.3</td>
</tr>
<tr>
<td>Item 23</td>
<td>1.44</td>
<td>-3.65</td>
</tr>
<tr>
<td>Item 24</td>
<td>1.3</td>
<td>-1.95</td>
</tr>
<tr>
<td>Item 25</td>
<td>0.91</td>
<td>-2.07</td>
</tr>
<tr>
<td>Item 27</td>
<td>1.86</td>
<td>-0.9</td>
</tr>
<tr>
<td>Item 28</td>
<td>0.92</td>
<td>-0.58</td>
</tr>
<tr>
<td>Item 29</td>
<td>1.12</td>
<td>-2.88</td>
</tr>
<tr>
<td>Item 31</td>
<td>2.01</td>
<td>-1.07</td>
</tr>
<tr>
<td>Item 32</td>
<td>1.15</td>
<td>-2.46</td>
</tr>
</tbody>
</table>

Note. Item parameters estimated using Bock-Aitkin EM estimation in IRTPRO. Wald tests comparing item parameters between the two groups are reported, all items have the same degrees of freedom. * $p < .05$. 
Table 6. Number of participants in each current religious affiliation by childhood religious affiliation.

<table>
<thead>
<tr>
<th>Current Religious Affiliation</th>
<th>Agnostic</th>
<th>Atheist</th>
<th>Baptist</th>
<th>Buddhism</th>
<th>Catholic</th>
<th>Fundamental</th>
<th>Islam</th>
<th>Judaism</th>
<th>Mormon</th>
<th>Protestant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agnostic</td>
<td>1815</td>
<td>83</td>
<td>336</td>
<td>118</td>
<td>1215</td>
<td>230</td>
<td>22</td>
<td>36</td>
<td>100</td>
<td>1030</td>
<td>4985</td>
</tr>
<tr>
<td>Atheist</td>
<td>252</td>
<td>188</td>
<td>104</td>
<td>19</td>
<td>311</td>
<td>52</td>
<td>10</td>
<td>11</td>
<td>17</td>
<td>285</td>
<td>1249</td>
</tr>
<tr>
<td>Baptist</td>
<td>10</td>
<td>1</td>
<td>345</td>
<td>2</td>
<td>15</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>32</td>
<td>423</td>
</tr>
<tr>
<td>Buddhism</td>
<td>98</td>
<td>13</td>
<td>25</td>
<td>221</td>
<td>112</td>
<td>21</td>
<td>3</td>
<td>4</td>
<td>10</td>
<td>102</td>
<td>609</td>
</tr>
<tr>
<td>Catholic</td>
<td>33</td>
<td>2</td>
<td>12</td>
<td>9</td>
<td>1017</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>42</td>
<td>1123</td>
</tr>
<tr>
<td>Fundamental</td>
<td>39</td>
<td>5</td>
<td>37</td>
<td>12</td>
<td>34</td>
<td>453</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>62</td>
<td>648</td>
</tr>
<tr>
<td>Islam</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>145</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>158</td>
</tr>
<tr>
<td>Judaism</td>
<td>11</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>13</td>
<td>5</td>
<td>0</td>
<td>86</td>
<td>0</td>
<td>13</td>
<td>133</td>
</tr>
<tr>
<td>Mormon</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>235</td>
<td>8</td>
<td>255</td>
</tr>
<tr>
<td>Other</td>
<td>243</td>
<td>20</td>
<td>162</td>
<td>24</td>
<td>386</td>
<td>84</td>
<td>17</td>
<td>10</td>
<td>27</td>
<td>431</td>
<td>1404</td>
</tr>
<tr>
<td>Protestant</td>
<td>168</td>
<td>7</td>
<td>298</td>
<td>20</td>
<td>138</td>
<td>152</td>
<td>0</td>
<td>3</td>
<td>14</td>
<td>1918</td>
<td>2718</td>
</tr>
<tr>
<td>Total</td>
<td>2676</td>
<td>319</td>
<td>1326</td>
<td>425</td>
<td>3250</td>
<td>1024</td>
<td>197</td>
<td>153</td>
<td>411</td>
<td>3924</td>
<td>13705</td>
</tr>
</tbody>
</table>
Table 7. Internal consistency reliability across ethnic groups, Study 2.

<table>
<thead>
<tr>
<th>Ethnic Group</th>
<th>Anti-Science</th>
<th>Pro-Magic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native American</td>
<td>.84</td>
<td>.80</td>
</tr>
<tr>
<td>African American</td>
<td>.83</td>
<td>.76</td>
</tr>
<tr>
<td>Chinese</td>
<td>.78</td>
<td>.73</td>
</tr>
<tr>
<td>Filipino</td>
<td>.74</td>
<td>.74</td>
</tr>
<tr>
<td>Hispanic</td>
<td>.81</td>
<td>.77</td>
</tr>
<tr>
<td>Indian</td>
<td>.81</td>
<td>.72</td>
</tr>
<tr>
<td>Japanese</td>
<td>.76</td>
<td>.72</td>
</tr>
<tr>
<td>Korean</td>
<td>.82</td>
<td>.73</td>
</tr>
<tr>
<td>Middle Eastern</td>
<td>.85</td>
<td>.79</td>
</tr>
<tr>
<td>Pacific Islander</td>
<td>.82</td>
<td>.78</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>.76</td>
<td>.66</td>
</tr>
<tr>
<td>Caucasian</td>
<td>.86</td>
<td>.79</td>
</tr>
</tbody>
</table>
Figures

Figure 1. Bidimensional representation of nMyst subscales as quadrants.

- **High - Low:** Uses only AS to protect their beliefs. Reject paranormal as evil, likely fundamental religiosity.
- **Low - Low:** Does not have a strong need for any mysterious beliefs. Materialists.
- **High – High:** Uses both strategies to protect their beliefs. Most successful, likely very common.
- **Low - High:** Uses only PM to protect their beliefs. No objection to science.
Figure 2. CFA Final Model with standardized estimates, Study 2.

Note. CFA run in EQS using robust ML estimation, $\chi^2(177) = 361.89$, $p < .001$, $CFI = .95$, $TLI = .93$, $RMSEA = .05$ [.04 - .05]. Primary factor loadings are in bold, red lines indicate the covariance is set to one. * Estimated parameter.
Figure 3. Comparison of total information curves for European Heritage and Chinese Heritage samples, Study 2.

Note. Information is a measure of the accuracy of the scale at each level of $\theta$ and is a function of the standard error of measurement at each level. The Chinese heritage sample has greater information than the European heritage sample at most $\theta$ levels for both AS and PM subscales.
Figure 4. Means of Anti-Science comparing non-converts to converts to and from each religion.

Note. Comparison between non-converts and converts-from is significant, $F(1, 13652) = 146.29$, $p < .001$. Only Agnosticism, Buddhism, and Judaism are not significantly different between non-converts and converts-from. Comparison between non-converts and converts-to is not significant, $F(1, 12484) = .02$, $p = .90$. Differences between non-converts and convert to Atheism and Buddhism were significant, $F(1, 12484) = 16.15$, $p < .001$ and $F(1, 13652) = 10.22$, $p < .001$, respectively.
Figure 5. Means of Pro-Magic comparing non-converts to converts to and from each religion.

Note. Comparison between non-converts and converts-from is significant, $F(1, 13652) = 18.20, p < .001$. Only Islam and Judaism showed no significant difference between non-converts and converts-from. Comparison between non-converts and converts-to is not significant, $F(1, 13652) = 1.63, p = .20$. Atheists and Mormon showed a significant difference between non-converts and converts-to, $F(1, 12484) = 10.17, p < .001$ and $F(1, 12484) = 5.07, p < .05$ respectively.
Figure 6. Ethnicity ANOVA results Study 2.

Note. ANOVA is significant for AS, $F(12, 14406) = 6.39, p < .001$, and PM, $F(12, 14406) = 8.48, p < .001$. Sample sizes: Native American ($N = 161$), African American ($N = 376$), Chinese ($N = 1669$), Filipino ($N = 343$), Hispanic ($N = 717$), Indian ($N = 199$), Japanese ($N = 97$), Korean ($N = 307$), Middle Eastern ($N = 114$), Pacific Islander ($N = 61$), Vietnamese ($N = 157$), Caucasian ($N = 9860$).
Figure 7. ANOVA results of education level by Anti-Science and Pro-Magic, Study 4.

Note. Effect is significant for both AS and PM, $F(5, 14595) = 24.88, p < .001$ and $F(5, 14595) = 24.56, p < .001$, respectively.
References


316(5827), 996–997. doi:10.1126/science.1133398


doi:10.1177/0146167211399583


studies? A comparative analysis of six preconceptions about Internet questionnaires.

*American Psychologist, 59*, 93-104.


Nemeroff, C., & Rozin, P. (2000). The makings of the magical mind: The nature and function of sympathetic magical thinking. *Imagining the impossible: Magical, scientific, and


Appendix A

Initial 32 item set from Study 1. Responses were given on a 5 point Likert scale from Strongly Agree to Strongly Disagree. Items retained for Study 2 are italicized.

1. I prefer not to know how magic tricks are done.
2. When reading or watching a mystery movie, I need to know the ending as soon as possible.
3. Learning about Darwin's evolutionary theory doesn't interest me.
4. Love is just a chemical reaction.
5. There are some things that humans should not try to understand.
6. I am comforted by the idea of a higher power that loves us.
7. I find it hard to believe in things that I can't picture.
8. There are things in this world too complicated to happen on their own.
9. Reports of so-called 'death experiences' don't prove there is an afterlife.
10. I prefer spiritual explanations for things rather than scientific ones.
11. Given that so many people are religious, there must be something to it.
12. If someone told me they had a spiritual vision I would doubt it.
13. I believe there is something beyond this material universe.
14. Some people try to make the world too complicated.
15. I sometimes think that inanimate things are fighting against me (e.g., computers, cars).
16. I never think of objects as having a personality.
17. Animals have souls too.
18. Objects sometimes move around on their own.
19. It is cruel to cut down trees.

20. Some houses are definitely haunted with ghosts.

21. I am comfortable with the idea that the world is just atoms and molecules.

22. I trust my intuition more than scientific evidence.

23. I follow my instincts, even if I don't understand them.

24. I sometimes get a shiver thinking about mysterious things.

25. Looking up at the stars thrills me.

26. I sometimes get a shiver hearing certain musical pieces.

27. When I'm on my own, I sometimes get magical feelings.

28. I sometimes get goosebumps when I'm really connecting with people.

29. I often get the feeling that I've done something before - even though it's not possible (deja vu).

30. I check my horoscope regularly.

31. Some people can predict the future.

32. I tend to believe in certain conspiracy theories.
Appendix B

Paranormal belief items from Study 1.

1. Hypnosis
2. Extrasensory Perception
3. Telekinesis
4. Psychic/ Channeling
5. Astrology
6. Tarot
7. Ghosts
8. Exorcism
9. Vampires
10. The Lost City of Atlantis
11. The US government organized 9/11
12. The Freemasons control the world
13. Aliens
14. Alien abduction
15. The US government has aliens at Area 51
16. UFOs
17. Prayer can affect things in the world
Appendix C

Final 20 item set after dropping a poorly performing item in Study 2.

Anti-Science:

1. There are some things that humans should not try to understand.
2. There are things in this world too complicated to happen on their own.
3. Reports of so-called "near-death experiences" don't prove there is an afterlife.
4. I prefer spiritual explanations for things rather than scientific ones.
5. If someone told me they had a spiritual vision I would doubt it.
6. I believe there is something beyond this material universe.
7. I am comfortable with the idea that the world is just atoms and molecules.
8. I trust my intuition more than scientific evidence.
9. I follow my instincts, even if I don't understand them.

Pro-Magic:

10. I sometimes think that inanimate things are fighting against me (e.g., computers, cars).
11. I never think of objects as having a personality.
12. Animals have souls too.
13. Objects sometimes move around on their own.
14. Some houses are definitely haunted with ghosts.
15. I sometimes get a shiver thinking about mysterious things.
16. When I'm on my own, I sometimes get magical feelings.
17. I sometimes get goose-bumps when I'm really connecting with people.
18. I check my horoscope regularly.
19. Some people can predict the future.

20. I believe in some conspiracy theories.
Appendix D

Items used from the Schizotypal Personality Questionnaire (Raine, 1991) and Hood’s Mysticism Scale (1975).

SPQ Items

Ideas of reference items:

1. Do you sometimes feel that things you see on the TV or read in the newspaper have a special meaning for you?
2. I am aware that people notice me when I go out for a meal or to see a film.
3. Do some people drop hints about you or say things with a double meaning?
4. Have you ever noticed a common event or object that seemed to be a special sign for you?
5. Do you sometimes see special meanings in advertisements, shop windows, or in the way things are arranged around you?
6. When shopping do you get the feeling that other people are taking notice of you?
7. When you see people talking to each other, do you often wonder if they are talking about you?
8. Do you sometimes feel that other people are watching you?
9. Do you sometimes feel that people are talking about you?

Magical thinking and paranormal belief items:

1. Have you had experiences with the supernatural?
2. Do you believe in telepathy (mind-reading)?
3. Are you sometimes sure that other people can tell what you are thinking?
4. Do you believe in clairvoyancy (psychic forces, fortune telling)?
5. Can other people feel your feelings when they are not there?

6. Have you had experiences with astrology, seeing the future, UFOs, ESP, or a sixth sense?

7. Have you ever felt that you are communicating with another person telepathically (by mind-reading)?

Unusual perceptual experiences items:

1. Have you often mistaken objects or shadows for people, or noises for voices.

2. Have you ever had the sense that some person or force is around you, even though you cannot see anyone?

3. When you look at a person, or yourself in a mirror, have you ever seen the face change right before your eyes?

4. I often hear a voice speaking my thoughts aloud.

5. Have you ever seen things that are invisible to other people?

6. Do everyday things seem unusually large or small?

7. Does your sense of smell sometimes become unusually strong?

8. Do you ever suddenly feel distracted by distant sounds that you are not normally aware of?

9. Are your thoughts sometimes so strong that you can almost hear them?

**Mysticism Items**

**General mysticism:**

1. I have had an experience which was both timeless and spaceless.

2. I have had an experience in which everything seemed to disappear from my mind until I was conscious only of a void.
3. I have never had an experience in which I felt myself to be absorbed as one with all things. R

4. I have never experienced a perfectly peaceful state. R

5. I have never had an experience in which all things seemed to be aware. R

6. I have had an experience in which I realized the oneness of myself with all things.

7. I have had an experience in which I felt that all was perfection at that time.

8. I have had an experience in which I felt everything in the world to be part of the same whole.

9. I have had an experience that is impossible to communicate.

10. I have never had an experience in which my own self seemed to merge into something greater. R

11. I have never had an experience in which time, place, and distance were meaningless. R

12. I have had an experience in which all things seemed to be conscious.

13. I have had an experience that cannot be expressed in words.

Religious mysticism:

1. I have had an experience in which something greater than myself seemed to absorb me.

2. I have experienced profound joy.

3. I have never had an experience which seemed holy to me. R

4. I have never experienced anything to be divine. R

5. I have never experienced anything that I could call ultimate reality. R

6. I have had an experience which I knew to be sacred.
7. I have had an experience which left me with a feeling of awe.

8. I have never had an experience which left me with a feeling of wonder. $R$

9. I have never had an experience in which deeper aspects of reality were revealed to me. $R$