RICE UNIVERSITY

Measuring and Predicting Extreme Response Style: A Latent Class Approach

by

Bobby Darius Naemi

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

Master of Arts
Psychology

APPROVED, THESIS COMMITTEE:

Daniel Beal, Assistant Professor, Chair
Margaret Beier, Assistant Professor
Michelle “Mikki” Hebl, Associate Professor

HOUSTON, TEXAS
MAY 2006
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.
ABSTRACT

Measuring and Predicting Extreme Response Style:
A Latent Class Approach
by
Bobby Naemi

The purpose of this study was to explore various ways to predict and measure extreme response style, or overuse of endpoint categories in rating scales. Data was collected from a total of 913 regular participants and 240 peer participants, who completed an online battery of self-report and peer report questionnaires respectively. In addition to verifying the stability and generality of extreme responding, extreme response style was related to two personality predictors: intolerance of ambiguity and decisiveness. Both main effects and interactive effects with speed of survey completion were uncovered. Extreme response style was measured with several simple proportional methods, which were all shown to tap a latent factor of response extremity, and a latent class method, which did not achieve significant relationships with the personality predictors.
Acknowledgements

I would like to thank my advisor, Dan Beal, for his invaluable support, patience and guidance. I would also like to thank my committee members, Margaret Beier and Mikki Hebl, for their helpful input and flexibility.
<table>
<thead>
<tr>
<th>Table of Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
</tr>
<tr>
<td>Introduction</td>
</tr>
<tr>
<td><em>Response Styles and Response Sets</em></td>
</tr>
<tr>
<td><em>Predicting Extreme Response Style</em></td>
</tr>
<tr>
<td><em>Intolerance of Ambiguity</em></td>
</tr>
<tr>
<td><em>Decisiveness</em></td>
</tr>
<tr>
<td><em>Revising Measures of Intolerance of Ambiguity and Decisiveness</em></td>
</tr>
<tr>
<td><em>Hypothesis 1a, 1b, 2a</em></td>
</tr>
<tr>
<td><em>Hypothesis 2b, 2c, 3a, 3b</em></td>
</tr>
<tr>
<td><em>Hypothesis 4, Interpretational Benefits of Peer Reports</em></td>
</tr>
<tr>
<td><em>Measuring Extreme Response Style</em></td>
</tr>
<tr>
<td><em>Hypothesis 5</em></td>
</tr>
<tr>
<td><em>Mixed Rasch Model</em></td>
</tr>
<tr>
<td><em>Hypothesis 6</em></td>
</tr>
<tr>
<td><em>Hypothesis 7</em></td>
</tr>
<tr>
<td>Method</td>
</tr>
<tr>
<td><em>Participants</em></td>
</tr>
<tr>
<td><em>Procedure, Materials</em></td>
</tr>
<tr>
<td><em>Extreme Response Style Measure</em></td>
</tr>
<tr>
<td><em>Intolerance of Ambiguity Measure, Decisiveness Measure</em></td>
</tr>
<tr>
<td><em>Time</em></td>
</tr>
<tr>
<td>Results</td>
</tr>
</tbody>
</table>
Measurement Equivalence 36
Reactions to Ambiguity 39
Decisiveness 40
Extreme Response Style 40
Hypothesis 1: Convergent Validity, 42
Hypothesis 2: IA and ERS 43
Hypothesis 3: Decisiveness and ERS 44
Hypothesis 4: Time, Decisiveness and ERS 45
Hypothesis 5: ERS Measures, Hypotheses 6-7: Mixed Rasch Analyses 46
Discussion 48
References 60
Tables 70
Figures 75
Appendix 81
List of Tables

Table 1 List of All Personality Measures in Questionnaire Battery, by Section

Table 2 Means, Standard Deviations, and Correlations between Demographic and Personality Predictors of Extreme Response Style (TAMU Sample, Self and Peer Reports Only)

Table 3 Factor Loadings for TAMU vs. Rice Participants

Table 4 Differing Factor Models of the Reactions to Ambiguity Scale

Table 5 Consistent Akaike's Information Criterion Values Determining Number of Classes
Measuring and Predicting Extreme Response Style: A Latent Class Approach

Surveys are an extremely popular method of choice for measurement in all fields of psychology. In industrial and organizational psychology in particular, surveys enjoy widespread use in areas such as performance appraisal, supervisor ratings, and assessment of employee perspectives and attitudes. Approximately 70% of small, midsized, and large companies in America conduct employee surveys, and 69% report plans to conduct further organizational surveys in the future (Gallup, 1988). Among large corporations, the prevalence of organizational survey use is approximately 78% (Kraut, 1996) and both the number of companies who use organizational surveys and the frequency of organizational survey use are increasing (Borg, 2003). Use of psychological surveys is clearly ubiquitous, and researchers and employers alike must rely on accurate information from these surveys when making workplace decisions. It is thus crucial to ensure that the methods of measurement employed in psychological surveys produce information that is both valid and reliable, given the potential consequences of employment or workplace decisions derived from survey information. Promoting accurate and consistent methodological techniques is thus an integral part of workplace research, in terms of both employing measures that are fair, and ensuring that measures are useful.

Although psychological surveys can contain many different types of measurement methods, most surveys make use of some sort of ordinal response or Likert scale. The overwhelming frequency of Likert scale use highlights the need to examine any potential confounds that may result from this choice of measurement method. Given that inferences of individual attitudes, traits and behaviors are generally drawn from Likert
scale responses, researchers have long been interested in how features of the Likert scale itself may influence category response choices. Research has shown that the number of response categories (Givon & Shapira, 1984; Preston & Coleman, 2000), presence of a midpoint or neutral response option (Gilljam & Granberg, 1993; Klopfer & Madden, 1978) and the verbal labels of anchor points (French-Lazovik & Gibson, 1984; Lam & Klockars, 1982; Schaeffer, 1991) can all influence a person's response on a Likert-type scale.

Response Styles and Response Sets

Research has also examined how factors associated with the respondents themselves influence category choices on Likert scales. In particular, both response style and response set, terms that are frequently confused in the literature, can affect Likert scale category choices. I shall use the terminology specified by Rorer (1965), who defined response set as a pattern of responses that is dependent on the content or format of items, and response style as a consistent pattern of responses that occurs irrespective of item content and format. Although it is possible that response styles might occur only some of the time, in theory, response styles should occur consistently across all items.

Examples of response set include social desirability, the tendency to give favorable or advantageous responses, and deviance, the tendency to give deceptive responses. A respondent will thus employ these response sets for some items or measures but not others, depending upon whether the content elicits a motivation to appear socially desirable or deviant.

In contrast, response styles reflect a pattern of responding that is consistent across item formats and content. Examples of response style include acquiescence, the tendency
to agree or say "yes" when responding (also known as leniency, or yea-saying), standard deviation, the tendency to use a narrow or wide range of responses on a scale (also known as response range), and extreme response style, the tendency to disproportionately favor the endpoints or extreme categories of a scale, which is the focus of this paper.

Response styles also have been conceptualized as existing along a continuum. Specifically, the reverse of leniency has been defined as severity (or nay-saying) and the reverse of extreme response style has been defined as central tendency or the tendency to overuse the middle category of a scale.

When comparing response styles, extreme response style can be distinguished from leniency, the tendency to use only high endpoints and severity, the tendency to only use low endpoints, because extreme response style refers to the alternate use of both the high and low endpoints on a scale. Extreme response style and standard deviation are also often compared as similar response styles, and although they are highly correlated, they are not identical (Greenleaf, 1992a). Standard deviation takes into account response proportions on all points of a scale, whereas extreme response style typically only counts the proportion of responses of the endpoints of a scale.

Response styles and response sets thus both present potential problems for the researcher. However, a distinction can be made between response style and response set, in that response set occurs based on the content or features of the scale used in the measure. Both of these factors, the content of the measure and the format of the scales, can theoretically be manipulated by the researcher. Thus, if an experimenter wanted to reduce the effect of social desirability or deviance in a measure, he/she could change the wording of the items in order to discourage these response biases. Similarly, if an
experimenter wanted to reduce possible confounds as a result of features of the scale, he/she might be able to adjust the number of response categories, add/remove a midpoint or change the labels of the response categories. However, a researcher cannot influence response style, because it is generally impossible to adjust or change aspects of the respondents themselves. Thus, research involving response style may be particularly critical because it affects an area that is theoretically out of the experimenter’s control. Because attributes of the respondent may influence whether a response style is employed or not, it then becomes important to examine the possible individual difference antecedents of response style, which is also the focus of this study.

Researchers have argued that response styles act as potential confounds in accurate measurement, but it is also important to determine whether these possible effects are of any substantial consequence in terms of test validity and reliability. Nunnaly and Berstein (1994) posited that differing response styles between survey respondents can act as a nuisance variable that interferes with the measure of the variable of interest. This is especially problematic when seeking to compare individual responses to one another, given that differences in scale scores should only reflect differences in attitudes and behaviors, and not result from differences in response style. When this assumption, which is referred to as measurement invariance, is violated, mean differences between groups or individuals cannot be validly compared (Eid & Rauber, 2000). Indeed, Guion (1998) has suggested that understanding response styles and their effects on employee surveys and inventories “should be a high research priority” (Guion, 1998, p. 597) in industrial and organizational psychology.
The detrimental effects of response style on survey validity were highlighted by researchers as early as Cronbach (1946), who suggested certain methods for controlling and eliminating these effects on test validity. Research on response style was popular throughout the 1950’s and 1960’s until an influential critique by Rorer (1965) entitled “The Great Response Style Myth” was published in *Psychological Bulletin*. Rorer reviewed the literature regarding the influence of response styles on personality questionnaires and concluded that both response sets and response styles had a negligible effect on scale validity. Rorer’s article focused mainly on acquiescence, and despite rejoinders from other researchers (Rundquist, 1966), this critique is credited with the decline of research on response style that occurred after the 1970’s (Nunnally & Bernstein, 1994; Schimmack, Bockenholt & Reisenzein, 2002).

Despite acting as a seemingly devastating blow to response style research, Rorer’s (1965) article and other similar critiques made little or no reference to a response style that has admittedly received less research attention: extreme response style. Hamilton (1968) and Merrens (1970) pointed out that many of the criticisms of acquiescence made by Rorer and others do not generalize to extreme response style. Specifically, although Rorer found that acquiescence did not reliably emerge as a response style across measures that differed in content, research has shown that extreme response style occurs reliably across differing content measures (Greenleaf, 1992b; Hamilton, 1968; Jain & Agrawal, 1977; Merrens, 1970), potentially negating Rorer’s criticisms of the generality and stability of response styles (as described below).

In terms of concrete effects of response style differences on the validity and reliability of measures, researchers have recently returned to this question, focusing on
extreme response style in particular. Johnson (2003) found that failing to account for individual differences in extreme response style resulted in a tendency for the apparent effects of explanatory variables to become attenuated. Even more dramatically, Johnson demonstrated that conclusions regarding the relation between two factors can be completely different due to differences in extreme response style. Specifically, when a simple homogenous model of measurement was employed (a model that assumed all respondents used the scale identically), the correlation between two factors was positive (.19) but when a heterogeneous thresholds structure model was employed (a model that allowed for individual variation in scale use, or differing response styles), the correlation between the two factors was negative (-.25). Johnson concluded that the differences in response style (namely extreme response style) used by participants could be obscuring the true relation between the factors. Thus, simply using a model that allows for differing category thresholds for Likert scale responses can result in a more accurate assessment of the underlying relation between factors.

In another study, Eid and Rauber (2000) identified distinct clusters of individuals who differed in terms of in response style in a sample of 4578 participants, identifying a separable group of extreme responders. The researchers argued that comparing groups who differ in extreme response style violates measurement invariance, referring to the principle that groups cannot be meaningfully compared if they are not equivalent. The validity of experimental conclusions relating non-equivalent subgroups that differ in terms of response style can thus be called into question.

Finally, Greenleaf (1992a) found that standard deviation, a response style that is very similar to extreme response style, contained a bias component, or systematic
variation that occurred beyond basic attitude information from the scales. Results showed that if this bias component remained uncorrected, respondents were assigned to different group categories or attitude classifications than if scale scores were corrected for response style. Greenleaf thus argued that failing to correct for the bias component in response style could lead to a respondent being assigned inaccurately to a category that is not reflective of their true attitudes.

In short, several recent studies have shown that failing to account for extreme response style can result in confounded or spurious correlations, indicating the importance of extreme response style as a confounding variable in survey research. Having established that extreme response style can act as a potentially consequential nuisance variable in group comparisons based on survey research, it is important to confirm the stability and reliability of this particular response style. Contrary to Rorer’s (1965) critique of acquiescence as a response style, almost all research on extreme response style has shown that it is a stable tendency that occurs reliably across differing measures (Greenleaf, 1992b; Hamilton, 1968; Jain & Agrawal, 1977; Merrens, 1970), usually by demonstrating that extreme responding occurs across measures of varying content, or reliably over time. There are a few studies, however, indicating that extreme response style may be unstable in certain situations (Hui & Triandis, 1985; Innes, 1977). Hui and Triandis in particular found that extreme response style was more likely to occur in items at the end of a lengthy questionnaire, and attributed this result to potential fatigue or boredom on behalf of the respondents. Although Hui and Triandis did find this pattern occurred for two separate questionnaires on two different occasions, the researchers did not make a strong effort to control for the content of the scales, thereby making it
difficult to determine the extent to which similar content at the end of the questionnaire was responsible or if fatigue was indeed the appropriate explanation. Greenleaf proposed that other studies also indicating instability of extreme response style may have been improperly measuring extreme response style by failing to employ content-varied measures. In concordance with Greenleaf, I anticipate no differences in extreme response style; however, in the current study, I will attempt to account both for fatigue and content explanations for potential differences.

*Predicting Extreme Response Style*

Given that most research has shown extreme response style seems to be a stable trait that occurs reliably over time, it seems likely that there are certain stable personality traits that are characteristic of extreme responders. In short, extreme response style can be considered a behavioral manifestation of stable personality traits that differ among individuals. As such, it is beneficial to determine the dispositional antecedents of extreme response style, in order to better predict who is most likely to employ this confounding technique and when it is most likely to be used. In this sense, extreme response style can be argued to be an outcome variable that occurs as a result of certain personality characteristics which differ among individuals. Researchers have attempted to determine these dispositional antecedents by exploring the relations between extreme response style and various personality and demographic variables.

The strongest evidence relating extreme response style and demographic variables comes from cross-cultural research, in which researchers have found robust evidence of reliable cross-cultural differences in the use of extreme response style (Baumgartner & Steenkamp, 2001; Clarke, 2000; Gerardo et al, 1992; Hui & Triandis, 1989; Johnson et
al., 2005; Van Herk et al., 2004). Only one of these studies, however, offers any systematic explanation for these differences across various cultures. The particular study (Johnson et al., 2005) examined a sample of 19 nations and found that Hofstede’s (2001) cultural dimensions of masculinity and power distance were positively related to extreme response style. The implications of these findings on the relations between personality variables and extreme responding are discussed in more detail in the section on decisiveness. In short, findings from cross-cultural studies have led researchers to argue that comparisons across cultures must control for extreme response style, a concern that is particularly relevant for international organizations that rely on survey data from people of differing nationalities (Clarke, 2000; Johnson, 2003, Johnson et al., 2005).

Regarding other demographic variables, some studies have shown that women tend to exhibit more extreme response style than men (Berg & Collier, 1953; Borgatta & Glass, 1961; Crandall, 1973) although several studies have failed to find any gender differences (Brengelmann, 1960b; Greenleaf, 1992b; Light et al, 1965). Further research has found that Hispanics and African-Americans tend to exhibit more extreme response style than White Americans (Bachman & O’Malley, 1984; Gerardo et al, 1992), that children and adolescents tend to make more extreme responses than adults (Hamilton, 1968; Light et al, 1965; Soueif, 1958), that younger adults tend to make less extreme responses than older adults (Das & Dutta, 1969; Greenleaf, 1992b), that individuals of lower occupational status and education levels exhibit more extreme response style than those at higher levels (Eid & Rauber, 2000; Greenleaf, 1992b; Litinsky & Haslam, 1998; Soueif, 1958), and that extreme response style is present in mental health patients exhibiting depression or suicidal tendencies (Litinsky & Haslam, 1998; Teasdale et al.,
2001). Although these studies are interesting and informative, again, none offer a theoretical interpretation of why extreme response style might occur, other than the fact that there might be some form of cognitive ability component to extreme response style, given demographic differences in terms of age, education level and mental instability.

Interestingly, however, research conclusions from studies investigating the relation between extreme response style and intelligence have been mixed, in that some studies have established significant negative relations (Brengelmann, 1960a; Das & Dutta, 1969; Light et al., 1965; Wilkinson, 1970), whereas others have found none (Kerrick, 1954; Zuckerman & Norton, 1961). Research associating extreme response style and authoritarianism has been similarly mixed, with experimenters reporting significant positive correlations, significant negative correlations and non-significant correlations between extreme response style and authoritarianism measures such as the F-scale (see Hamilton, 1968 for a review).

In terms of other personality variables, studies have shown that extreme response style is related to anxiety (Berg & Collier, 1953; Lewis & Taylor, 1955) although it is unclear whether this is a direct relation or determined by other variables such as sex or intelligence (Crandall, 1973; Kerrick, 1954).

In short, most research relating extreme response style to personality and cognitive ability variables seems to be mixed, a result that may be attributed to Greenleaf's (1992a) contention that past research is plagued by inconsistencies in the measurement of extreme response style. Indeed, in a review of the past literature, Hamilton (1968) found that extreme response style did not relate significantly to any standard personality inventories. However, a number of older studies have managed to suggest a significant relation
between extreme response style and one particular dispositional variable: intolerance of ambiguity.

_Intolerance of Ambiguity_

The construct of intolerance of ambiguity was first discussed by Frenkel-Brunswik (1949) and was later formally defined by Budner as "the tendency to perceive ambiguous situations as sources of threat" (Budner, 1962, p. 29). Budner also defined an ambiguous situation as "one which cannot be adequately structured or categorized by the individual because of the lack of sufficient cues" (Budner, 1962, p. 30). These cues comprise the following three types of ambiguous situations, as defined by Budner: "a completely new situation in which there are no familiar cues, a complex situation in which there are a great number of cues to be taken into account, and a contradictory situation in which different elements or cues suggest different structures" (Budner, 1962, p. 30). The three types of ambiguous situations correspond to Budner's three components of ambiguity: novelty, complexity and insolubility. Although Budner's is certainly not the only definition of intolerance of ambiguity, (see Furnham & Ribchester, 1995, for a review) it is one of the earliest definitions that attempted to provide a theoretical coherence and clarity to the otherwise poorly defined construct of intolerance of ambiguity.

Intolerance of ambiguity represents a compelling potential predictor of extreme responding because unlike past research involving extreme response style and demographic variables, studies relating intolerance of ambiguity and extreme responding suggest an underlying theoretical mechanism for exhibiting extreme response style. Individuals with a high intolerance of ambiguity should theoretically avoid making equivocal responses and instead gravitate towards definite unambiguous choices. As
such, high intolerance for ambiguity could serve as a mechanism to explain why a respondent would avoid using potentially ambiguous middle categories on a rating scale, and instead make disproportionate use of the definite unambiguous endpoints of a rating scale. Given the research indicating that extreme response style seems to be a stable trait that differs between individuals, it seems likely that a stable personality construct such as intolerance of ambiguity may act as an explanatory individual difference variable for extreme responding.

In terms of past literature relating intolerance of ambiguity and extreme response style, Soueif (1955) suggested that extreme response style could serve as an indicator of intolerance of ambiguity in a series of studies, although he provided no support for his contention. Brim and Hoff (1957) also found that intolerance of ambiguity was significantly correlated with extreme response style on three out of four measures (correlations ranging from .28-.57). Finally, Brengelmann (1960a) found consistent significant correlations between an intolerance of ambiguity measure and extreme responses in nine different samples (correlations ranging from .28-.45), and suggested that extreme response style was definitively associated with the so-called rigid personality type. Based on this research, Brengelmann (1960b) concluded that extreme response style can be seen as a way of achieving structure and order in the testing measure, thus reducing ambiguity.

Despite the promise shown in the previous research, past studies relating intolerance of ambiguity and extreme responding are unfortunately plagued by serious problems relating to accurate construct measurement that call into question whether it is truly intolerance of ambiguity that predicts extreme responding. Specifically, Brim & Hoff’s
(1957) study employed a measure of need for certainty as a proxy for intolerance of ambiguity, and Brengelmann's (1960a) measure of intolerance of ambiguity was in actuality primarily a measure of social dominance, consisting mainly of items such as "Life is a battle in which the strongest get the prize." In addition, in Brengelmann's study, the relation between intolerance of ambiguity and extreme responding was only shown to occur for positive extreme responses, as opposed to true extreme responding on both endpoints of the scale. The methodological problems in this past research, which include a high degree of construct contamination and poor psychometric properties of measurement scales, make it difficult to tell whether it is in fact intolerance of ambiguity or other similar constructs such as social dominance, rigidity or need for certainty that predict extreme responding. As such, one purpose of this study will be to provide an improved assessment of the relation between extreme response style and intolerance of ambiguity by developing new measures for both these constructs.

The concept of intolerance of ambiguity has also recently been recognized as belonging to a cluster of related traits dealing with epistemic motives. Lay epistemic motives refer to stable dispositional motives relevant to how people think about themselves and others (Kruglanski, 1989). Thus, an intolerance of ambiguity, or cognitive rigidity, affects how an individual cognitively organizes any received information and how they apply this information to their decisions or responses to situations. In this context, intolerance of ambiguity represents an epistemic method of organizing information about ambiguous stimuli, and extreme response style represents a behavioral manifestation of this epistemic motive. As such, demonstrating a relation between extreme response style and intolerance of ambiguity can serve as a form of
validation of intolerance of ambiguity as an epistemic construct. In order to demonstrate this relation, however, it is imperative that uncontaminated, psychometrically sound measures of both constructs are employed.

In short, the first purpose of this study will be to confirm the relation between intolerance of ambiguity and extreme response style, using more valid and reliable questionnaire measures for these constructs (see below).

Decisiveness

Another personality predictor that may serve as a potential theoretical explanation of the mechanism behind extreme response style is the construct of decisiveness. Decisiveness has been globally defined as encompassing the ease, speed and confidence of making firm decisions (Kruglanski, 1989; Thompson, Naccarato, & Parker, 1989). Neuberg et al. (1997) proposed that the construct of decisiveness, or fear of invalidity, is associated with an initial desire for a nonspecific answer or closure in a situation. Similarly, decisiveness may drive an individual to employ an extreme or definite response in a questionnaire.

Theoretically, decisiveness represents a new and appealing personality predictor of extreme response style. A dispositional tendency to quickly seize on a solid answer that provides the strongest possible decision might lead respondents to commit to a definite extreme response, whereas indecisiveness may lead to respondents avoiding extreme categories that represent strong decisive opinions. This hypothesized relation has some precedence in cross-cultural research relating the cultural dimension of masculinity and extreme response style. Johnson et al. (2005) posited that features of masculine cultures that emphasize assertiveness and decisiveness might lead respondents to choose the
strongest available categories to express their opinions, and the researchers consequently found a significant relation between masculinity and extreme responding, contradicting past research indicating that women are more likely to use extreme responses. Although it is possible that other elements relevant to masculinity (including potentially intolerance of ambiguity) may have led to extreme response style, it is worth including decisiveness as a potential predictor of extreme responding based on the theoretical reasons highlighted above. As such, the personality construct of decisiveness may account for individual differences in extreme response style above and beyond differences accounted for by intolerance of ambiguity.

Items on decisiveness scales such as the decisiveness facet of the Need for Closure Scale (Kruglanski and Webster, 1994) include reverse-scored statements such as “I tend to struggle with most decisions” and “When trying to solve a problem I often see so many possible options that it's confusing.” These items seem to suggest that someone who is indecisive will struggle with the many possible choices available to them and not be able to make a quick confident decision. Such a person would be unlikely to employ an extreme response style when responding to Likert scale items on a survey, because a person who is able to see many possible choices and generally struggles when deciding on a response option would be likely to use the entire scale in a less ordered, uniform way. However, a person who is high on decisiveness might have no problems committing to a firm response and not agonize over response choices, instead making quick confident decisions. Such a person might be more likely to make an extreme or definite response, primarily using the endpoints of Likert scales as a way of confidently and reliably
validating their decisions. In short, decisiveness could serve as a potential personality antecedent of the decision to use an extreme response style.

*Revising Measures of Intolerance of Ambiguity and Decisiveness*

There are several modern self-report measures for the constructs of both intolerance of ambiguity and decisiveness. One popular measure, the Need for Closure Scale, assesses the degree to which a person seeks "*an* answer on a given topic, *any* answer... compared to confusion and ambiguity" (Kruglanski, 1990, p. 337, italics in original). The process of achieving closure is divided into two distinct steps, an initial "seizing" of an answer, relating to a person's urgent desire to obtain any closure as quickly as possible, and a subsequent "freezing" on that answer, reflecting a person's desire to maintain and protect the previously achieved closure (Webster & Kruglanski, 1996). The NFCS consists of 44 items divided into five facets: Preference for Order, Preference for Predictability, Decisiveness, Discomfort with Ambiguity and Close-mindedness. The researchers argue that each of these facets can contribute to the construct of need for closure in differing ways. The two facets of discomfort with ambiguity and decisiveness represent measures of the constructs of intolerance of ambiguity and decisiveness respectively.

The discomfort with ambiguity facet consists of items such as "I dislike it when a person's statement could mean many different things" and "I'd rather know bad news than stay in a state of uncertainty" (Kruglanski, & Webster, 1994). Kruglanski and Webster (1994) found that items belonging to the Discomfort with Ambiguity facet correlated moderately (.36) with Eysenck's (1954) Intolerance of Ambiguity Scale, and attributed this relatively low correlation to the psychometric failings of the older Eysenck measure.
The decisiveness facet consists of items such as "I tend to struggle with most decisions" and "When trying to solve a problem I often see so many possible options that it's confusing." Both of these items also appear on an alternate measure of decisiveness, Thomson et al.'s (1989) 14-item Personal Fear of Invalidity (PFI) scale. Personal Fear of Invalidity reflects the perceived costs of errors in decision-making, and thus serves as a measure of indecisiveness or cautiousness when making decisions (Thomson et al., 1989). Neuberg et al. (1997) found median correlations of -.75 between the NFCS facet of decisiveness and the oppositely coded PFI scale, and concluded that these two measures were essentially redundant.

In addition to using these previously validated measures of intolerance of ambiguity and decisiveness to measure the constructs, it is also beneficial to develop new more logically coherent self-report measures of these constructs in terms of how they relate to extreme response style. In terms of intolerance of ambiguity, although Budner's (1962) definition of ambiguity helped to provide a coherence and clarity to the construct lacking in previous treatment of the concept, Budner's measure of intolerance of ambiguity, a 16-item scale consisting of items assessing positive or negative perceptions, feelings and behaviors involved in ambiguous situations, possessed poor internal reliability (.49).

As such, it is useful to develop a new more coherent measure of intolerance of ambiguity, both on psychometric and conceptual grounds. Thus, a new "Reactions to Ambiguity" self-report questionnaire will be developed that incorporates items from previous scales purporting to measure intolerance of ambiguity (Budner, 1962; Norton, 1975; Webster & Kruglanski, 1994) as well as newly written items.
Budner (1962) categorized an ambiguous situation as one that is complex, contradictory, or novel. In developing a new definition of ambiguity, I will incorporate the first two components of complexity and contradiction, but I will argue that the third, novelty, is not a requirement for ambiguity in and of itself and should thus be discarded from definitions of ambiguity. Instead, the concept of novelty should be replaced with the component of incompleteness/vagueness, a concept that has appeared in Norton’s (1975) multidimensional Intolerance of Ambiguity scale.

The key to determining whether the components of complexity, contradiction and vagueness/incompleteness are truly defining attributes of ambiguity is to test whether these components, in and of themselves, can produce ambiguity. Applying this test to the first element of ambiguity, we can see that complexity, which is defined as the amount of information or number of cues associated with a stimulus (e.g. a complicated person with dozens of personality traits) in and of itself, is sufficient to create ambiguity, because having a large number of cues or a vast amount of information to process makes something difficult to understand or unclear.

Regarding the second element of ambiguity, we can also see that contradiction, which is defined as the presence of opposite or conflicting information (e.g. a person with both good and bad traits) in and of itself, is sufficient to create ambiguity, given that objects that contain conflicting information are more difficult to understand and not clear-cut and unambiguous.

Thus, complexity and contradiction seem to be elements that produce ambiguity, in and of themselves. However, when we apply this test to the element of novelty, we can see that the ambiguity of a situation does not arise because it is new and unfamiliar, but
rather because it is vague or incomplete. This is not merely semantic quibbling. The ambiguity of a new situation is not based solely on the pure novelty or unfamiliarity of the situation in and of itself. Instead this ambiguity is due to a lack of information, or lack of completeness. The reason a situation is ambiguous is not because it is simply new, it is ambiguous because it has not yet formed a definitive complete impression, resulting in too little information to process the situation. Thus, an item on an ambiguity measure such as “A person who leads an even, regular life with few surprises or unexpected happenings has a lot to be grateful for” (Budner, 1962) is not truly assessing ambiguity at all, but rather reactions to novelty or new situations. Preference for familiarity and dislike of unexpected new situations are concepts that are more closely tied to the constructs of routinization of habits, or assigning structure and order to life, and do not necessarily belong in the construct definition of intolerance of ambiguity. Because these items do not sufficiently capture the definitive qualities of a situation that are ambiguous, they should be discarded from questionnaire measures of intolerance of ambiguity, and may indeed have contributed to the poor psychometric performance of Budner’s own measure of intolerance of ambiguity.

Regarding the personality variable of decisiveness, a new self-report measure of decisiveness was also developed with the intention of focusing solely on the tendency to quickly “seize” on an answer, as opposed to including confidence or fear of making errors. Although alternate measures of decisiveness or indecisiveness such as the PFI scale and the NFCS decisiveness facet do not face the same poor psychometric performance problems of Budner’s (1962) Intolerance of Ambiguity scale and other intolerance of ambiguity measures (see Furnham & Ribchester, 1995), there is no
theoretical reason to believe that fear of making errors, or confidence in decision-making, might be related to extreme response style. However, the tendency to quickly seize and commit to the strongest possible answer may predict tendency to commit to a solid extreme response, given that an extreme response potentially represents the strongest choice on a rating scale. Despite the fact that both pre-existing measures are presumably unidimensional, it is useful to develop an alternate scale focusing solely on tendency to quickly seize on an answer alongside these established measures, in the hopes of ascertaining the most parsimonious explanation of how decisiveness might be related to extreme response style.

Having described the personality constructs and the reasoning for their proposed relations with extreme response style, we can now make the following formal hypotheses:

Both of the newly developed scales for intolerance of ambiguity and decisiveness will exhibit convergent validity with older measures.

H1a: The Discomfort with Ambiguity facet of the NFC scale and the newly developed Reactions to Ambiguity measure will exhibit convergent validity.

H1b: The Decisiveness facet of the NFC scale, the Personal Fear of Invalidity scale, and the newly developed Decisiveness scale will all exhibit convergent validity.

In addition, both intolerance of ambiguity and decisiveness will predict extreme response style above and beyond previous measures of these constructs, and past demographic predictors of extreme response style.

H2a: Intolerance of ambiguity, as measured by the Reactions to Ambiguity scale, will be positively related to extreme response style.
$H2b$: The Reactions to Ambiguity measure of intolerance of ambiguity will possess incremental validity over the existing measures of discomfort with ambiguity.

$H2c$: The Reactions to Ambiguity measure will possess incremental validity over gender, ethnicity, cognitive ability and decisiveness as predictors of extreme response style.

$H3a$: Decisiveness, as measured by the newly developed Decisiveness scale, will be positively related to extreme response style.

$H3b$: The Decisiveness measure will possess incremental validity over gender, ethnicity, cognitive ability and intolerance of ambiguity as predictors of extreme response style.

Any interactive effects of intolerance of ambiguity and decisiveness on extreme response style will also be explored, although there is no theoretical reason to expect an interaction.

Given that decisiveness, or tendency to seize on a strong solid answer, is hypothesized as being a predictor of extreme response style, it is also theoretically possible that the speed at which a survey is completed may influence likelihood of exhibiting extreme response style. In this sense, people who spend less time on the survey might be more likely to make extreme responses, simply because extreme responses are the simplest and least time consuming choice to make when differentiating among Likert scale categories (an extreme response is essentially the equivalent of YES or NO, or AGREE or DISAGREE). However, people might also speed through the survey because they are actually confident in their answers and find the survey easy to use, or the content of the survey may represent important or highly accessible information (e.g. Fazio, 1990).
As such, time spent on the survey may not predict extreme response style in and of itself, but people who quickly speed through the survey and are also highly decisive may be likely to exhibit extreme response style. Thus, even if time spent on the survey is not noticeable enough to emerge as a main effect, it is possible that people who complete a survey quickly and are high in decisiveness might be particularly likely to use an extreme response style.

_**H4:** The negative relation between time and extreme response style will be moderated by decisiveness, such that the relationship will become stronger as decisiveness increases (see Figure 1).

This hypothesis thus suggests that individuals who are high in decisiveness and complete the survey quickly will be the most likely to exhibit an extreme response style.

**Interpretational Benefits of Peer Reports**

One important detail to note is that because all of the previous measures of intolerance of ambiguity and decisiveness make use of Likert scales, these measures may themselves be confounded by the extreme responses of the respondents. In this sense, when attempting to correlate the self-report measures of intolerance of ambiguity and decisiveness with extreme response style, it is important to somehow account for the fact that the predictor measures are also potentially contaminated with extreme response style.

A possible way of addressing this problem is to use peer ratings of personality variables in lieu of self-report predictors. By asking peers who are familiar with the respondent to fill out each of the personality predictors as if they were the respondent, the potential confounds of the respondent's own extreme responses can be avoided. Research has indicated that peer ratings of personality constructs are both highly accurate and
reliable alternatives for self-reports, with peer ratings even occasionally serving as better behavioral predictors than self-report measures (Barbaranelli & Caprara, 2000; Barrick, Patton, & Haugland, 2000; Biesanz & West, 2000). Using peer ratings as alternatives to self-report ratings can thus potentially allow for a more accurate assessment of the relation between the predictor variables and extreme response style.

Measuring Extreme Response Style

Having specified the measures for intolerance of ambiguity and decisiveness, it is also important to develop an accurate measure of extreme response style. In this sense, the proposed study advances the past literature, which has measured extreme response style in strikingly different ways. Researchers have measured extreme responses by counting endpoint use in Likert scales of varying lengths, (Berg & Collier, 1953; Soueif, 1955) by counting only positive extreme responses (Brengelmann, 1960), by collapsing categories in longer Likert scales (Borgatta & Glass, 1954) or even using standard deviation as a measure of extreme responding (Hamilton, 1968). Given that several of these measures purporting to assess extreme response style appear to be measuring different response styles entirely, (leniency or acquiescence in the case of Brengelmann (1960), standard deviation in the case of Hamilton (1968)), providing a standardized method of assessing what constitutes an extreme response may add a welcome clarity to the past literature.

The proposed study will thus endeavor to measure extreme response style in many different ways, in order to ascertain whether these differing measures of extreme response style all contain an underlying core of response extremity.
The first measure of extreme response style will be a simple score consisting of the proportion of extreme response category use across all items in the questionnaire battery (except for the items used in other specific ERS measures, as described below). Using such a summative measure across all items in a battery is the most common method of measuring extreme response style in the literature (Hamilton, 1968). Because each item in the battery represents an opportunity for an extreme responder to make an extreme response, using all the items in the scale, coded for response extremity, can provide a very large scale to measure ERS behavior.

This approach has one potential flaw: because the items are all taken from scales that are designed to be highly correlated, then it is possible that an overall measure of extreme responding will be tainted by the multiple "pockets" of items that are highly related. Therefore, another measure of extreme response style is a so-called “contentless” measure designed by Greenleaf (1992) for the explicit purpose of measuring extreme response style. Greenleaf’s measure incorporates sixteen items from an extensive survey administered to a large US sample in 1975 and 1987. These items were selected from the questionnaire based on their low inter-item correlations and equivalent response category proportions, meaning that these items are essentially free of content influences and can be considered representative of a respondent’s scale use on any six category Likert scale. The proportion of extreme response scores on these sixteen items was shown to be stable over a twelve year period, further buttressing the argument that extreme response style is a stable trait.

Although Greenleaf’s measure of extreme response style represents an improvement over past measures due to its explicit removal of content influences, the items comprising
this measure were individually selected post hoc from various sections of a lengthy questionnaire. As such, presenting the sixteen items of the Greenleaf scale in only one location might result in differing inter-item correlations when compared to Greenleaf's original survey data. In order to address this potential problem, a third measure of extreme responding can be constructed using Greenleaf's technique on the new set of data. By selecting items from the current questionnaire battery that are shown to have low inter-item correlations and similar response proportions, any potential problems with presenting the Greenleaf items in one particular location of the questionnaire battery can be avoided.

A fourth method of measuring extreme response style is to include a measure of a purposefully ambiguous object. For example, respondents in organizations are often required to rate candidates on Likert scale measures either for performance evaluations or during the hiring process. If the true scores for such candidates fall in the middle of a scale, then the extent to which people use the extreme categories should reflect a very sensitive measure of extreme response style. For the purposes of this study, respondents were asked several questions about their willingness to hire a candidate for a hypothetical management position based on a sample resume that was purposely designed to be ambiguous (see Measures). Thus, asking candidates to rate a seemingly ambiguous, unexceptional resume represents a real world situation where extreme responders might differentiate themselves from non-extreme responders, who would presumably rate this ambiguous resume using the middle categories of the scale as opposed to an extreme endpoint.
Each of these methods of measuring extreme response style presumably captures response extremity in conceptually distinct ways. The overall measure of ERS has the largest number of items and thus represents the largest scale measure of ERS, but this measure fails to take into account inter-item correlations, skew and response proportions across categories. The Greenleaf measures of ERS do take into account these factors, and thus represent a smaller, more content-varied measure of response extremity across items chosen to be psychometrically equivalent. Finally, the resume rating scale represents a specific measure of extreme responding in a real-world context: reactions to an ambiguous stimulus object (in this case, a resume). Whether these differing conceptual approaches to measuring ERS result in significantly different measures of extreme response style is an empirical question, but it is expected that these four measures of ERS will reflect some general latent tendency for extreme responding. As such, a factor analysis of these four ERS measures can be conducted to determine whether there is an essential core to extreme responding that generalizes across differing types of scales in the questionnaire.

H5: Factor analysis of the four differing measures of extreme response style will result in a unidimensional construct of “response extremity” that generalizes across all measures of ERS.

Each of these measures of extreme response style share one thing in common, and that is the idea of using proportions to measure extreme response style on a continuum, such that higher scores indicate greater occurrence of extreme response style. However, a novel approach to measuring extreme response style is to assume that there is a distinct category of extreme responders that is separate from a class of normal responders,
meaning that extreme response style occurs as a dichotomy rather than a continuum. This approach has some precedence based on studies of political extremism (e.g. Van Hiel & Mervielde, 2003), where extremists are often hypothesized as being categorically distinct from moderates. Similarly, extreme responders may belong to a separate class of respondents than non-extreme responders. One modern statistical technique that might allow for this type of categorical measure of extreme response style is the use of the polytomous Mixed Rasch Model.

**Mixed Rasch Model**

The Mixed Rasch model is a type of latent class model that allows for respondents to be assigned to separate classes that emerge from the data, based on their response styles. MIRA models are hybrid latent-trait, latent class models that essentially uncover subgroups or subpopulations within a data set based on patterns of responses (McCollam, 1998). The polytomous Mixed Rasch model is defined as (Rost, 1991):

\[ P_{ig}(y = k|\theta) = \exp(k\theta + e_{ikg})/\sum_{k'} \exp(k'|\theta + e_{i(k'|g)}), \]

where \( k \) is response category, \( k' \) is the next highest response category, \( i \) is item, \( g \) is latent class, \( \theta \) is an individual parameter of the respondent’s trait level, \( e_{ikg} \) is a parameter of response category difficulty, and \( \exp \) is the transcendental number \( e \). This equation thus represents the probability (\( P_{ig}(y=) \)) of a respondent in a latent class \( g \) choosing the response category \( k \) of item \( i \), given the person’s trait level \( \theta \) and the difficulty of choosing the particular response category, or category threshold difference \( e_{ikg} \) (Rost, 1997).
The Mixed Rasch model can thus be used to estimate two distinct latent classes of respondents, one comprised of regular respondents with ordered category thresholds, and another comprised of extreme respondents with representatively extreme category threshold differences. Specifically, if a class is identified containing large threshold differences between the endpoint response categories, this indicates that respondents in this class find it very difficult to cross the threshold of choosing an extreme category. Such a respondent requires very high or very low trait scores in order to use the highest or lowest category in a rating scale and is thus a typical, non-extreme responder. If, however, another class is identified containing relatively small threshold differences between the endpoint response categories, this indicates that respondents in this class find it easy to cross the threshold of choosing an extreme category, meaning that respondents may use the endpoints on rating scales even if their underlying trait scores are not exceedingly high or low. Such a class of respondents can clearly be defined as extreme responders, indicating that they have a tendency to use the extreme categories on scales, even when their underlying trait scores do not warrant the use of such categories.

Applying the Mixed Rasch model to any of the scales in the battery can thus allow for extreme respondents to be identified by examining the probability that the respondent will choose an endpoint category, holding their underlying trait level (which reflects the content of the measure) constant. Given that the Mixed Rasch model can identify extreme responders through adaptive cut-off scores that are customized to each individual person’s data, this type of latent class analysis represents a new, more sophisticated measure of the individual differences between “extreme” and “non-extreme” respondents. The Mixed Rasch Model also provides particular benefits over other techniques used to
identify homogenous subgroups (e.g. discriminant function analysis, cluster analysis) because MIRA analyses employ latent trait scores as opposed to observed variable scores that do not incorporate models of item responding (Zickar, Gibby & Robie, 2004). The distinction in this case is crucial, since a cluster analysis would likely assign respondents to an extreme response category based on their mean observed score across items, whereas a MIRA analysis identifies classes by examining the pattern of responses across items, holding latent trait level constant. A MIRA analysis takes into account not only whether a certain number of items in a scale were endorsed, but which particular items were endorsed, when assigning respondents to classes. Thus, two respondents may have the same trait scores on a measure of extraversion, for example, but could be assigned to different classes based on item parameter estimates. A cluster analysis is unable to distinguish between someone exhibiting an extreme response style across extraversion items, and someone who is actually high in extraversion, and would thus assign both respondent types to an extreme responder class.

Because a MIRA analysis can detect whether two respondents with the same trait-level respond differentially to the same item, MIRA analysis also serves as a way of detecting item bias in terms of differential item functioning (DIF). DIF refers to when participants who belong to different subgroups respond differently to the same item despite holding trait-level constant. Using a MIRA analysis thus not only helps prove that extreme response style can lead to the test bias of differential item functioning, but also more accurately distinguishes between respondents using an extreme response style and respondents with truly extreme trait scores.
It is in this capacity that the true worth of using Mixed Rasch modeling to measure extreme response style becomes apparent. Past measures of extreme responding have relied almost exclusively on simple continuous proportions or counts of endpoint use to measure extreme response style. A simple count of number of extreme categories used says nothing about whether the extreme category was chosen because of the bias of response style, which represents a potential confound to data, or whether the extreme category was chosen as an actual reflection of extreme attitude or belief, which represents a legitimate data point. Employing MIRA analyses that can distinguish between whether respondents are truly high in trait score or simply biased towards extreme categories thus represents an important advance in the extreme response style literature. There is also some precedence for using latent class modeling to identify extreme responders, as classes of extreme respondents have been shown to emerge in several studies (Austin et al., 1998; Eid & Rauber, 2000; Johnson, 2003; Rost, Carstensen, & von Davier, 1997; Zickar, Gibby & Robie, 2004).

By differentiating separate classes of respondents based on category threshold differences, the MIRA analysis will thus do a better job of getting at what it is to be an extreme responder, and in this way, result in better correlations with the personality predictors of intolerance of ambiguity and decisiveness. This leads to the following hypotheses:

$H_6$: A MIRA analysis will indicate that a two class model has best fit, such that two main classes of respondents will emerge, one consisting of extreme responders, and one of non-extreme responders.
H7: The correlation between the personality predictors of intolerance of ambiguity and decisiveness and extreme response style will be significantly stronger using the MIRA-based measure of ERS than the proportion-based measure of ERS.

By examining measurement equivalence, a better understanding of the psychometric differences between extreme and non-extreme responders can be reached.

Given that this study is also identifying new ways of measuring extreme response style using proportion methods (through the alternate Greenleaf, resume-rating and overall ERS rating scales), it is important to verify that extreme responding occurs reliably and stably in this particular set of data. By using so many different measures of extreme responding across a wide array of content items, the generality or reliability of extreme responding across different questionnaire items can be confirmed (see Methods).

In addition, the stability of extreme responding throughout the questionnaire can also be tested based on these data. Hui & Triandis (1985) proposed that extreme response style was more likely to occur at the end of a questionnaire due to fatigue or boredom with the measure, which seemingly counteracts the idea that extreme response style is a primarily stable dispositional trait that is characteristic of a particular profile of extreme responders. Thus it is predicted that in this questionnaire measure, the order in which the scales are presented will have no effect on likelihood of extreme responding. It should be noted, however, that although we expect no main effect of scale order on extreme responding, if such an effect does occur, it will be possible to assess whether it is a result of the content of measure (indicating a lack of content invariance for extreme response style) or time spent completing the survey (indicating fatigue, as hypothesized by Hui & Triandis, 1985).
In short, it is proposed that extreme response style is a primarily dispositional stable variable that is reflective of certain stable personality traits that differ among individuals, namely intolerance of ambiguity and decisiveness. The following study will also attempt to determine the best way of measuring extreme response style in light of its relation with these personality predictors. This study will thus contribute to the past literature on extreme response style in the following ways: verifying the relation between extreme response style and intolerance of ambiguity using improved measures of these constructs; establishing a new relation between extreme response style, decisiveness, and speed of survey completion; examining measurement differences in past proportion measures of extreme response style; determining the extent to which scale order influences extreme responding in lengthy questionnaires; and finally, developing a new psychometrically sophisticated measure of extreme response style through Mixed Rasch modeling.

Method

Participants:

Data was collected from three separate study samples. The first study sample consisted of 522 undergraduates at Texas A&M University, 132 of whom also provided peer rating data. The second study sample consists of 127 students at Rice University (the purpose of this sample was primarily to provide a sufficient sample size for IRT analyses, and thus no peer ratings were collected). The third study sample consisted of 233 undergraduates in the summer program at Texas A&M University, 62 of whom provided peer rating data.
Procedure:

Participants completed a battery of self-report questionnaire measures in an online survey over approximately a one-hour period, in exchange for course credit. Participants were also encouraged to recruit peer respondents who knew them well to complete all of the same questionnaire measures as if they were the participant. Peer respondents were instructed as follows: "Please answer all questions AS IF YOU WERE YOUR FRIEND/THE PARTICIPANT OF THE EXPERIMENT. Do NOT answer the questions with your own personal responses. If you feel you do not know how your friend might respond to a particular question, or you feel that you don’t have enough information to answer a question, please take a guess as to how you feel your friend might respond." In order to provide an incentive for the peer respondents, two $50 prizes were offered to two randomly selected peer respondents.

Materials:

The questionnaire battery consisted of three sections containing 21 separate measures (see Table 1) as well as the sample resume. The questionnaire was divided into three sections for the purposes of varying the content location. Three separate versions of the questionnaire were then created using the following orders:

Version 1 – Section 1, Section 2, Section 3

Version 2 – Section 2, Section 3, Section 1

Version 3 – Section 3, Section 1, Section 2.

Respondents in the initial TAMU sample were randomly assigned to one of these three versions based on the last 4 digits of their social security number, resulting in relatively equal proportions of respondents for each version of the questionnaire. The following
sections provide details of the specific measures included in the questionnaire used to assess the constructs of interest (i.e., Extreme Response Style, Intolerance of Ambiguity, and Decisiveness).

*Extreme Response Style:*

The first measure of ERS used was a simple score consisting of the proportion of extreme response category use across all 289 items in the questionnaire battery (excluding items in the Greenleaf and resume measures of ERS).

The second measure of ERS used was Greenleaf’s (1992) 16-item “contentless” scale.

The third measure of ERS used was an alternate version of the Greenleaf scale, with sixteen new items selected from across the entire 273-item questionnaire battery based on their low inter-item correlations and similar response proportions across categories.

The fourth measure of ERS used was the 10-item resume rating scale, which assessed reactions to a purposefully ambiguous job resume (see Appendix). The resume was designed to be ambiguous by containing information that was contradictory (e.g. high GPA at an unranked business school, low GPA at a prestigious university), complex (e.g. numerous job positions in locations across the country) and vague (e.g. mention of “office skills” as opposed to specific program knowledge). Extreme response style was thus measured by obtaining the proportion of responses in the endpoints of the Likert scale for the ten questions assessing the ambiguous resume and candidate.

A final measure of ERS used was the probability of membership in the extreme respondent class, a value derived from MIRA analyses of the scales (see Results for the details behind which scales were used to calculate this value).
**Intolerance of Ambiguity:**

A 32-item self-report questionnaire of intolerance of ambiguity was developed that incorporated items from previous scales purporting to measure intolerance of ambiguity (Budner, 1964; Norton, 1975; Webster & Kruglanski, 1994) as well as newly written items. The questions were designed to tap reactions to three theoretically defined attributes of ambiguity: complexity, contradiction and vagueness. Items were written or chosen to reflect discomfort or negative reactions to complexity, contradiction and vagueness. Thus, the scale was created to be maximally flexible, allowing for an overall score of Intolerance of Ambiguity as well as facet scores for Intolerance of Complexity, Intolerance of Contradiction, and Intolerance of Vagueness. All items were written such that higher scores indicate greater intolerance of ambiguity (see Appendix). In addition, a previously used measure of intolerance of ambiguity, the 8-item discomfort with ambiguity facet of the Need for Closure Scale (Webster & Kruglanski, 1994) was used for validation purposes.

**Decisiveness:**

A new 10-item measure of decisiveness was developed that focused on tendency to quickly “seize” on an answer, as opposed to including confidence or fear of making errors. This 10-item measure incorporates some items from the Personal Fear of Invalidity scale (Thomson et al., 1989) and the NFCS decisiveness facet (Webster & Kruglanski, 1994), and was written such that higher scores indicate greater decisiveness (see Appendix). In addition, both the 14-item Personal Fear of Invalidity scale (Thomson et al., 1989) and the 7-item NFCS decisiveness facet (Webster & Kruglanski, 1994) were used for validation purposes.
Time:

The online survey service provided a measure of time taken to complete the survey. This measure of time was calculated from data indicating the exact time each respondent both started and finished the online survey. As is the case with most time data, a set of outliers completed the survey over a much longer period of time (i.e., 6 – 12 hours), indicating that some respondents left the browser window open and returned to the survey hours later. Consistent with other time-based studies (see Ratcliff, 1993) times greater than three standard deviations from the mean were dropped for the purposes of this analysis. In addition, there is a possibility that some of the remaining participants with relatively high time measures may not have spent the entire time completing the questionnaire. Given that the respondents were not observed while completing the questionnaire, it is not possible to determine whether the respondents spent the entire listed time completing the questionnaire. For example, participants may have taken a break from responding to the questionnaire items, worked on another task, received a phone call, etc. As such, high time scores may not portray actual time spent answering items as accurately as low time scores.

Results

See Table 2. for means, standard deviations, and correlations between all predictors and measures of extreme response style. Reliabilities of all relevant measures are also reported on the diagonal of the matrix.

Measurement Equivalence:

Before any hypotheses were tested, measurement equivalence between the TAMU, Summer and Rice samples was assessed in order to determine whether these data could
be combined into a single set. Typical tests of measurement equivalence begin with an
examination of configural equivalence (whether the best-fitting factor structure is the
same across groups), followed by factorial equivalence (an empirical test of whether the
factor loadings are the same for each item across groups), scalar equivalence (an
empirical test of whether the intercepts are the same for each item across groups) and
structural equivalence (homology, or equivalent relations with other identified variables
across groups) (see Vandenberg and Lance, 2000, for a review of these terms).

Given the fact that the Reactions to Ambiguity scale was a newly developed,
unvalidated measure, measurement equivalence was initially assessed with an exploratory
approach. Specifically, exploratory factor analyses were conducted for all three samples,
and the number of factors extracted was determined using parallel analysis (Hayton et al.,
2004). This allowed for the comparison of the number of viable factors in each of the
samples. Immediately, it was clear that different solutions were appropriate for each
sample, although some consistencies were observed. For example, although three factors
were extracted in both the Rice and TAMU samples, significant discrepancies were
observed, along with a handful of consistencies. Careful inspection of Table 3. highlights
this point.

Ultimately, measurement equivalence was impossible to obtain using the full
Reactions to Ambiguity scale. Measurement equivalence was much better when smaller
clusters of items were used (see results of Reactions to Ambiguity measure), but given
the post-hoc nature of these clusters, the decision was made to use the breadth of the full
scale at the expense of measurement equivalence.
Similar problems establishing measurement invariance were discovered between the TAMU peer and Summer peer samples. As such, the initial 522-participant TAMU sample was chosen as the primary data set for analysis and hypothesis testing, due to its interpretable factor structure for the Reactions to Ambiguity scale, and its large sample size for both peer and self-reports (see Appendix for sample output from the Rice and Summer samples indicating mean differences on most measures, and lack of structural equivalence between hypothesized variables).

Furthermore, in order to ensure parity of responses for both the self-report and peer report groups, results were reported only for those participants who had both self and peer ratings. Although both self-report and peer report results of personality measures are included, it is important to note that peer report results are of primary importance in terms of interpretation. Because self-reported ratings of personality are confounded by the respondent's own extreme response style, they do not provide as valid or reliable assessments of personality as peer ratings. As such, peer ratings are of most importance when making conclusions, and self-report ratings are reported primarily for purposes of comparison. Finally, participants who failed to follow directions and responded carelessly to a reverse-scored item on the resume rating scale, or failed to respond to a majority of items, were eliminated from analyses. Failure to follow directions was assessed using the resume rating scale, which contained 10 items, one of which was reverse-scored. Nine items on the resume rating scale were used to predict the score of the one reverse-scored item. The residuals of this analysis provided an estimate of how unlikely the response on the reverse-scored item was given the responses on the other nine items. The results of
this process, and an examination of the data to remove participants with large portions of missing data, identified 16 people who were eliminated from subsequent analyses.

All hypotheses were thus tested with the resulting sample of 116 peer and self-report ratings. In terms of demographics, the resulting sample was 61% female, with ages ranging from 17-22 ($M = 19.1, SD = 1.0$). In terms of ethnicity, the sample contained 100 White participants, 1 African-American participant, 11 Hispanic participants, 3 Asian participants and 1 participant who marked “other.”

Reactions to Ambiguity: Initial exploratory factor analyses and reliability analyses for the three hypothesized facets resulted in 10-item complexity subscale, a 5-item contradiction subscale, and a 10-item vagueness subscale. This three factor model of reactions to ambiguity was tested against a one factor model for the Reactions to Ambiguity scale. The three-factor model was found to have significantly better fit (see Table 4) than a one-factor model of intolerance of ambiguity, despite high correlations between the three factors. A three-factor model with only four-items in each subscale was found to have the best fit.

For the purposes of the overall hypothesis tests, the three facets were combined into an overall 25-item measure that achieved a reliability of .91. Given that this study was primarily concerned with developing a reliable measure that could serve as a predictor of extreme responding, an overall measure was deemed sufficient for the purposes of hypothesis testing. Because the scale with the best fit would have reduced the number of items to only four items per scale, in the interests of protecting bandwidth of the predictor space of intolerance of ambiguity in comparison with the criterion space of extreme
response style (see Wittman & Sub, 2001), the full 25-item Reactions to Ambiguity
measure was used.

In addition, based on the correlation matrix, reactions to complexity, contradiction
and vagueness are very highly correlated, and also substantially correlated with their
responding peer reports, but not so much that they are identical. In this sense, the peer
reports of intolerance of ambiguity seem to be providing extra information or variance
than self-report ratings.

Self-reported values of reactions to ambiguity ($M = 3.90, SD = .65$) were slightly
lower than peer-reported values of reactions to ambiguity ($M = 4.05, SD = .75$).

*Decisiveness*: Initial exploratory and reliability analyses resulted in the elimination
of two items from the scale. The resulting 8-item scale was confirmed to be
unidimensional, and achieved a reliability of .80. Once again, self-reports of decisiveness
are highly correlated with peer reports (.540) but these ratings are not identical, indicating
that peer reports are potentially providing additional information.

Self-reported values of decisiveness ($M = 3.36, SD = .78$) were slightly lower than
peer-reported values of decisiveness ($M = 3.59, SD = .73$).

*Extreme Response Style*: To compute the first measure of ERS, the proportion of
endpoint use across all 273 items in the questionnaire battery was calculated, excluding
the items that already appeared in the Greenleaf, alternate Greenleaf, and resume rating
measures of extreme responding. A reliability analysis showed that the 273-item scale
had a reliability of .89 before even being recoded into extreme responses, indicating a
potential lack of content variance, although a factor-analysis across all items indicated
there were 76 factors based on a scree plot. The reliability of this measure after items
were coded as extreme responses (1 for any endpoint category use, 0 for anything else) was .97.

The second measure of ERS, the Greenleaf Scale, was shown to have a reliability of .67 even before being recoded into extreme responses, (.71 after coding for extreme responses) indicating a similar lack of content variance across the sixteen items. This result was supported by examining the inter-item correlations, which were mainly in the .2-.3 range, including a few interitem correlations as high as .38.

To compute the third measure of ERS, the alternate version of the Greenleaf scale, sixteen items were selected from across the entire 289-item questionnaire battery (excluding the sixteen original Greenleaf items and resume rating items). These items were deliberately chosen based on their low (<.1) correlations with one another, and similar response proportions across categories (to avoid highly skewed items). The reliability of the Alternate Greenleaf scale was -.02 before recoding into extreme responses (.67 after recoding) and relatively few interitem correlations were above .1.

Regarding the fourth measure of ERS, the 10-item resume rating scale, a factor analysis confirmed that the 10-item scale was a unidimensional measure, with 10 items achieving a reliability of .81 before recoding into extreme responses, and .89 after recoding into extreme responses.

The ERS measures are all substantially correlated with each other, with the exception of the resume rating measure. This is most likely due to the fact that the resume to be rated was not, in fact, ambiguous, as indicated by a highly skewed mean rating (4.89 out of 6). To test this, the mean resume rating value was rescaled so that the middle of the scale was set equal to zero. After running a simple regression model, the intercept
value in a model with no predictors was shown to be significantly greater than zero. As such, the mean resume rating was shown to be much higher than the expected mean rating of 3.5 on a 6 point scale, indicating that the resume was not rated as unexceptional or ambiguous, on average. Consequently, ERS scores for this single scale were much more dependent on participants’ attitudes regarding the resume, as opposed to a biased tendency for response extremity.

Exploratory factor analyses indicated that although the first three measures of ERS all loaded highly onto a single factor, the resume rating ERS measure had a much smaller loading (.4). A confirmatory model was tested in which the first three ERS measures were assigned to a single ERS factor, and this model was shown to have excellent fit (CFI = .994, chi-square = 12.88, df = 2).

As such, a single unit-weighted ERS score was calculated based on the mean of the ERS scores of the overall questionnaire battery, the Greenleaf scale, and the Alternate Greenleaf scale. This unit-weighted ERS score was used as the primary measure of ERS for subsequent hypotheses. The ERS unit-weighted score had a mean of .22, and a standard deviation of .13, indicating that on average, participants made extreme responses approximately 22% of the time.

*Hypothesis 1: Convergent Validity*

The intolerance of ambiguity facet of the Need for Closure Scale (Webster & Kruglanski, 1994) is correlated .623 with the overall 25-item Reactions to Ambiguity scale, a combined measure including the complexity, contradiction and vagueness scales, indicating good convergent validity.
Similarly, the seizing scale is highly correlated (.640) with the decisiveness facet of the Need for Closure Scale (Webster & Kruglanski, 1994) and highly negatively correlated (-.621) with the Personal Fear of Invalidity scale (Thomson et al, 1989) which measures indecisiveness.

The overall Reactions to Ambiguity scale and the Decisiveness scale also seem to be moderately correlated (.287).

_Hypothesis 2: IA and ERS_

_H2a:_ By examining the correlation matrix, it is clear that in general, the factors of the Reactions to Ambiguity Scale are positively correlated with the various measures of ERS.

For self-report ratings of reactions to ambiguity, only reactions to vagueness appears to have a small relationship with the measures of ERS (.23 overall). Peer report ratings of reactions to complexity, contradiction and vagueness on the other hand, appear to be moderately correlated with ERS (.30 range, .35 overall). This might indicate that peer reports of intolerance of ambiguity, because they are unaffected by the potential social desirability and extreme responses of the participants themselves, are a more accurate assessment of intolerance of ambiguity and the true relationship between intolerance of ambiguity and ERS.

Although these relationships hold for the first three measures of ERS, a different pattern emerges for the resume rating measure of ERS, as contradiction and vagueness seem to be substantially correlated with resume ratings for self-reports (.23 and .29 respectively), but not for peer reports. This provides further evidence that the resume rating measure of ERS is clearly different than the other measures of ERS.
\textit{H2b}: The Reactions to Ambiguity measure of intolerance of ambiguity possesses incremental validity over the existing measures of discomfort with ambiguity, but only for peer ratings. Specifically, for self-reports, the discomfort with ambiguity facet of the NFCS accounted for 9.8\% of the variance in ERS, and the Reactions to Ambiguity scale failed to account for any significant additional variance. However, for peer reports, the discomfort with ambiguity facet of the NFCS accounted for 8.2\% of the variance in ERS, and the Reactions to Ambiguity Scale accounted for an additional 5.2\% of the variance in ERS.

\textit{H2c}: The unique contribution of reactions to ambiguity, over and above gender, ethnicity (coded white vs. nonwhite), time (with outliers removed), self-reported SAT score, and decisiveness, was assessed. The results indicated that for self-reports, none of the predictors were significant, except decisiveness (R squared change = .073). However, for peer ratings, decisiveness accounted for 8.1\% of the variance in ERS, and reactions to ambiguity accounted for an additional 13.7\% of the variance in ERS.

\textit{Hypothesis 3: Decisiveness and ERS}

\textit{H3a}: By examining the correlation matrix, decisiveness appears to have a small relation with ERS for both self-report and peer-report ratings. Self-report ratings of decisiveness, as measured by the newly developed Decisiveness scale, were positively related to extreme response style (.24 with unit-weighted ERS score) and peer report ratings were also positively related to ERS (.29 with unit-weighted ERS score).

\textit{H3b}: The unique contribution of decisiveness, above gender, ethnicity, time, cognitive ability and intolerance of ambiguity as predictors of extreme response style was assessed for both self and peer reports. For self reports, decisiveness accounted for 4.6\%
of the variance in ERS, wiping out the effect of reactions to ambiguity by itself (4.8% of the variance). However, for peer reports, decisiveness accounts for an additional 6.8% of the variance in ERS, above and beyond the 15% of the variance accounted for by intolerance of ambiguity.

For self-reports, reactions to ambiguity and decisiveness accounted for 7.8% of the variance simultaneously ($\beta = .17, p = .08$ for ambiguity, $\beta = .19, p = .05$ for decisiveness). For peer-reports, reactions to ambiguity and decisiveness accounted for 17.4% of the variance simultaneously ($\beta = .31, p < .001$ for ambiguity, $\beta = .24, p < .01$ for decisiveness).

Having established significant relationships for both intolerance of ambiguity and decisiveness with extreme response style, the possible interactive effects of these predictors were also assessed. Significant interactions were found for both self-report ($\beta = .44, p < .001$) and peer reports ($\beta = .29, p < .001$). See Figures 2 and 3 for a plot of these interactions. Specifically, as decisiveness increases, participants who are intolerant of ambiguity are more likely to exhibit an extreme response style, whereas people who are tolerant of ambiguity are less likely to exhibit an extreme response style.

**Hypothesis 4: Time, Decisiveness and ERS**

A moderator regression analysis was carried out in order to determine whether decisiveness and time spent on the survey interact to predict extreme response style. The measures used were self and peer ratings on the Decisiveness scale, the unit-weighted ERS score, and the measure of time taken to complete the survey provided by the online survey service. A significant interaction was found for self-report ratings ($\beta = -.21, p = .03$), but not for peer report ratings of decisiveness ($\beta = -.13, p = .185$). After plotting
both interactions however, (see Figures 4 and 5), results indicated that participants who completed the survey quickly and were high in decisiveness were the most likely to exhibit an extreme response style for both self-reports and peer reports. Based on problems inherent in the time measure, it is important to focus more on the results for low time, as the results for high time are not as valid or reliable.

Hypothesis 5: ERS measures

As stated above, factor analysis of the four differing measures of extreme response style resulted in a unidimensional construct of “response extremity” that generalizes across the first three measures of ERS, but not the resume rating measure of ERS.

Hypothesis 6-7: Mixed Rasch Analyses

The following series of decision rules were applied to determine which of the 17 possible scales (excluding the predictor variables of intolerance of ambiguity and decisiveness) could be used to derive an ERS score based on Mixed Rasch analyses. These rules were based on general guidelines for using latent class models in personality measures (Rost & von Davier, 1997).

1) The measure must be unidimensional.

2) The measure must have a sufficient number of responses made in each category of the rating scale, such that all categories of the scale are used.

3) The measure must contain no more than two classes of respondents (an extreme class and a non-extreme class).

4) The proportion of respondents in the extreme class should be in line with that of past research (.10-.40 of the sample).
In order to assess whether an identified class of respondents consisted of extreme responders, graphs of category threshold parameters for each item of the scale were examined. When the category threshold parameter distances for the endpoint categories were very close to the nearest threshold in one class, and far apart in the other, the former class was identified as consisting of extreme responders (see Figures 6-11).

Although there were several unidimensional scales with sufficient numbers of categories used to examine latent classes, the hypothesized classes of one extreme and one non-extreme class only emerged for three scales: Conscientiousness, Need for Structure, and Need for Cognition. For many of the other usable scales, additional latent classes that emerged tended to reflect people who used the scales in a disordered fashion, rather than in an extreme fashion. These three measures were the only ones where a two-class model fit the data best (see Table 5) and the resulting two classes could be identified as an extreme and non-extreme class, based on category threshold differences (see Figures 6-11 for threshold parameters of the non-extreme and extreme classes on the Conscientiousness, Need for Structure and Need for Cognition scales).

Having identified an extreme class of responders for these scales, three ERS measures were calculated based on the value provided by WINMIRA indicating the probability of membership in the extreme class. These ERS scores were only slightly correlated with one another (see Table 2), and moderately correlated with the unit-weighted ERS score based on the three simple proportional measures.

Neither the self-report nor peer report Reactions to Ambiguity measures were correlated with the Mixed Rasch measure of ERS. Similarly, neither self report nor peer reports of Decisiveness were correlated with the Mixed Rasch measure of ERS, except in
the case of the ERS measure for Need for Structure. For this particular measure, self-reported seizing was slightly correlated with ERS scores (.21) and peer-reported seizing was moderately correlated with ERS scores (.28).

Finally, the stability of extreme response style across the survey battery was assessed by examining whether the order in which the questionnaire scales were presented affected the likelihood of extreme responding. ERS scores were calculated for the Greenleaf Scale, Self-Esteem Scale (Rosenberg, 1965), and Resume Rating scale, all of which appear in Section 3 of the questionnaire battery (see Table 1). Section 3 appeared either at the beginning, middle, or end of the questionnaire battery based on the particular version of the questionnaire battery given to the 522-person TAMU sample. Relatively even proportions of participants in this sample completed the three different versions, such that 32% completed Version 1, 35% completed Version 2, and 33% completed Version 3. Results indicated that version of the questionnaire, or whether the ERS scales appeared near the beginning, middle or end of the battery, had no significant effect on likelihood of extreme response style.

Discussion

Overall, these research findings provide relatively good support for the hypotheses dealing with predicting extreme response style, and mixed support for those dealing with measuring extreme response style.

Firstly, measurement equivalence between the TAMU, Rice and Summer samples could not be established. It is not surprising that the sample of undergraduates from a selective private university such as Rice were not equivalent to a sample of a large state school population such as Texas A&M University, given potential underlying differences
in cognitive ability (as evidenced by SAT scores) and academic motivation. Other potential dissimilarities include dispositional differences between those who chose to attend a relatively apolitical school in a large city, and those who chose to attend school in a highly conservative small town. These differences may have affected how participants interpreted the construct of intolerance of ambiguity in terms of social desirability, since Rice participants may have been less comfortable endorsing items that indicated discomfort with complexity or contradictions than TAMU participants. This interpretation is somewhat supported by evidence comparing the mean values of Reactions to Ambiguity across samples, which indicate that the average score for the TAMU sample \((M = 3.90)\) is significantly higher than the average score for the Rice sample \((M = 3.54)\) \((p < .05)\).

A surprising result, however, was the inability to establish measurement equivalence between the TAMU summer sample and the regular TAMU sample. It is possible that there are also underlying differences between students who choose to stay for summer school and those who completed regular academic semesters that may have resulted in a lack of measurement invariance. Perhaps summer students perceived intolerance of ambiguity in a different way than regular-year students because of differences in their motivation to stay in school, or similar social desirability problems as described in the case of Rice students (although there are no significant differences between the mean ambiguity scores for TAMU \((M = 3.90)\) and TAMU Summer \((M = 3.89)\)). In any case, it is clear based on these results, that the Rice, TAMU, and TAMU summer samples are not equivalent.
In addition, it is possible that TAMU summer peer responses were not equivalent to TAMU regular peer responses because students during the summer session were unable to attract peers who knew them extremely well to fill out accurate assessments of their personality (e.g. their closest friends may not have been taking summer classes). Finally, the summer peer sample size of only 63 students may have also contributed to the lack of measurement equivalence as a result of lack of power.

In terms of the predictor measures used in this study, the newly developed Reactions to Ambiguity Scale was shown to have good reliability as an overall measure (.91) despite the fact that it is undoubtedly multidimensional. The three hypothesized facets of reactions to complexity, contradiction and vagueness did provide significantly better fit than a one-factor model, although the fit was still not satisfactory (CFI = .80).

The hypothesized facets, by focusing specifically on positive reactions to simplicity, consistency and definitiveness, and negative reactions to complexity, contradiction and vagueness, provide a greater conceptual clarity than past measures of intolerance of ambiguity that incorporate a myriad of similar constructs such as fear of uncertainty, fear of novelty, routinization, resistance to change and close-mindedness. Despite the improved conceptual clarity, however, the items chosen to measure intolerance of ambiguity did not fit well with the conceptual model of reactions to complexity, contradiction and vagueness. These results may have been due to the limited number of items initially used to sample the respective domains, given that typical scale validation efforts involve writing considerably more items than will be used in the final scale measure. In addition, combining items assessing discomfort with ambiguous attributes (e.g. complexity, contradiction and vagueness) and preference for unambiguous attributes
(e.g. simplicity, consistency and definitiveness or completeness) may have contributed to
the lack of good fit, an interpretation supported by research affirming the separable nature
of positive affective and negative affective reactions (Cacciopo and Berntson, 1994).

Empirically, a model employing subscales of complexity, contradiction and
vagueness would provide close to satisfactory fit if they were reduced to four-items each
(CFI = .93). For the purposes of this study, however, the decision was made to keep the
longer versions of these scales, as creating a shorter version of an already pared down
construct definition would narrow the construct space to such a degree that it would lose
value as a predictor of extreme response style. By maintaining full-length scales in lieu of
pared-down measures, it is thus possible to achieve greater Brunswick symmetry between
the constructs of intolerance of ambiguity and extreme response style (Wittmann & Sub,
1999).

The overall Reactions to Ambiguity measure was shown not only to have
convergent validity with past measures of intolerance of ambiguity (specifically the
discomfort with ambiguity facet of the NFCS), but also to predict additional variance in
extreme responding above and beyond this past measure of ambiguity intolerance.
Although this result was only shown for the peer-report measures as opposed to the self-
report measures, this may have been due to the fact that the self-report measures were not
as reliable or accurate as the peer ratings. Specifically, social desirability may have
played a role in coloring participants own responses as to whether they preferred
complexity vs. simplicity, for example.

Similarly, the Decisiveness measure was shown to have convergent validity with
past measures, and because decisiveness has not yet been related to extreme response
style, no claims were made regarding whether the newly developed Decisiveness scale could predict variance in extreme responding above and beyond older measures.

In terms of demographic predictors of extreme response style, past studies had indicated mixed support for the possibility of gender and cognitive ability predicting ERS, and some support for ethnicity predicting ERS. This study found no relationship between gender or cognitive ability (although self-reported SAT scores may not have been a valid or reliable measure of cognitive ability) and ERS, and also no relationship between ethnicity and ERS. The lack of findings for ethnicity as a predictor of ERS may reflect a true lack of ethnic differences in ERS, but considering the limited sample of non-white participants, this conclusion is only tentative.

Past studies had indicated a possible positive relationship between intolerance of ambiguity and extreme response style, and this study found that a measure specifically devoted to reactions to attributes of ambiguity (and not fear of uncertainty, rigidity, closemindedness, resistance to change, routinisation, social dominance, or a host of other related constructs conflated in past measures of IA) accounted for a significant amount of variance in extreme responding. This relationship held after accounting for other demographic and personality predictors of extreme response style, and was particularly strong for peer reports as opposed to self-reports of the construct. As such, we now have a better idea of the extent to which this specific construct is related to extreme responding.

One novel contribution of this study was the uncovering of a significant relationship between decisiveness and extreme response style. Specifically, a tendency to quickly seize on the strongest possible answer was shown to account for a significant amount of
variance in extreme responding, above and beyond past predictors of extreme response style. This relationship was also shown to occur above and beyond simple speed of completion of the survey. Notably, a significant interaction was discovered, such that for those participants who finished the survey quickly, those who were high in decisiveness were far more likely to exhibit an extreme response style than those who were low in decisiveness. As such, simply speeding through a questionnaire is not sufficient to lead to extreme response style; one must also be highly decisive. This result makes sense, because it is possible for those who speed through a survey to choose to respond randomly, or respond with a central tendency, in order to quickly complete the survey. These results, however, lend support to the idea that if someone is highly decisive, then they will employ an extreme response style in order to quickly complete the survey.

It is unclear though, based on the plotted interactions of decisiveness and time on ERS, why the relationship between time and ERS actually increases for those who are indecisive, when one might expect that individuals who are indecisive and spend a long time completing the survey would actually be the least likely to exhibit an extreme response style. This result may be due to imperfections in the measure of time spent to complete the survey. Because participants completed this survey entirely online, it is possible that participants may have left the survey in the middle of completion and returned to it at a later time, thus leading to misleading results for those who had high survey completion times. Although efforts were made to remove extreme outliers, it is still possible that results for high completion times could be confounded by this problem, whereas the results for low completion times could not be confounded in such a fashion. As such, when interpreting this figure, it is important to place more value on the
differential results for low time, or those who completed the survey quickly, since the values for high time are not as valid.

In terms of additional interactions, one unexpected result was an interactive effect of intolerance of ambiguity and decisiveness on extreme response style. Specifically, based on the peer report data, as decisiveness increases, those who are intolerant of ambiguity are more likely to engage in extreme response style, whereas those who are tolerant of ambiguity are less likely to engage in extreme response style. This result is actually complemented by the results of the previously described interactive effects of time and decisiveness on extreme response style. Specifically, it is possible that people generally use response styles as a kind of heuristic or shortcut method of completing the survey as quickly as possible, a prediction supported by the principles of lay epistemic theory, which posit individual differences in the desire to achieve closure (Kruglanski, 1989). In this way, decisiveness, or tendency to quickly seize on solutions, may act as a proxy for need to use any response style, whereas intolerance of ambiguity may determine which specific response style is used to quickly complete the survey. As such, participants who are high in decisiveness and high in intolerance of ambiguity may use an unambiguous, extreme response style, whereas participants who are high in decisiveness and high in tolerance of ambiguity may use a central tendency response style, as these people are comfortable expressing themselves with middle categories. Needless to say, this interaction represents an intriguing possibility regarding the difference between personality predictors of use and choice of response style that merits further study.

In short, this study has found that both the personality traits of intolerance of ambiguity and decisiveness can predict tendency to use the endpoint categories in a
survey, lending support to the idea that extreme response style may be in part a
manifestation of dispositional traits.

In terms of the measurement of extreme response style, this study found that four
different measures of extreme response style including an overall endpoint count, a
contentless measure, an alternate contentless measure and ratings of a presumably
ambiguous resume, all tap an underlying latent construct of response extremity. Although
the resume rating measure did not tap into this latent construct to the same degree as the
other three measures, it is probable that this is because the resume rating was biased by
the content of the measure, given that the resume was not in fact ambiguous, (as
evidenced by the high mean rating 4.89 on a six-point scale, when a truly ambiguous
resume should have received an average rating of 3.5).

The results of the simple proportional measures of ERS lend support to the idea that
there is a latent construct of extreme response style that generalizes across difference
measures, but this idea was not supported by the results of the Mixed Rasch analyses. It
was hypothesized that using a Mixed Rasch approach would more accurately distinguish
between responders who consistently choose endpoint categories because of response
bias, and those who make extreme responses because of high underlying trait levels.
Unfortunately, the hypothesized latent classes of extreme and non-extreme responders
only emerged in three out of the seventeen possible measures used in the questionnaire
battery. In addition, where these classes did emerge as hypothesized, they were not
significantly related to the personality predictors of intolerance of ambiguity or seizing,
except in the case of the ERS measure derived from the Need for Structure, which was
moderately related to seizing. It is unclear why the hypothesized relation only emerged for this particular measure.

In addition, the Mixed Rasch measures of ERS were only slightly related to one another, indicating that the participants identified as extreme responders with this method were not the same across the different scales. The Mixed Rasch measures were, however, moderately related to the overall unit-weighted ERS score, possibly indicating that they were tapping at least some part of a latent ERS construct.

One reason the Mixed Rasch method did not accurately identify extreme responders in this questionnaire battery may have had to do with the content of the particular measures employed. Specifically, there could have been some confounding between the content of the measures and the probability of membership in the extreme responder category. Because we anticipated that high trait levels in many of the personality measures used would lead to extreme responding, the Mixed Rasch method may have ruled out respondents who should have been classified as extreme responders, since respondents with legitimately high trait levels are disqualified from membership in the ERS category of a Mixed Rasch analysis. This interpretation, however, cannot account for the lack of accurate identification of extreme responders for the Self-Esteem Scale (Rosenberg, 1965) or the Agreeableness Scale, two personality measures that had no relationship with extreme response style.

A more likely explanation of why the Mixed Rasch measures did not accurately identify extreme responders is because of the relatively small sample size of 492 participants, once all respondents with partially missing data were eliminated. Typical analyses involving mixed-IRT models employ sample sizes in the thousands, so the
relatively small sample size may have reduced the accuracy of identification of extreme responders.

Finally, the order in which extreme response style measures were presented had no significant effect on likelihood of extreme responding, indicating a lack of evidence for the fatigue hypothesis advanced by Hui & Triandis (1985). As such, these results suggest that extreme response style is not unstable or dependent on length of a questionnaire, lending support to the idea that extreme response style can be predicted by stable dispositional traits.

In sum, this study has found good support for the idea that significant portions of the variance in extreme response style can be accounted for by two dispositional traits, intolerance of ambiguity, and decisiveness. Future studies should examine whether other personality variables can account for additional unique variance in extreme responding. In terms of predicting extreme response style, several interesting interactive effects were also uncovered that imply that non-dispositional variables such as speed of survey completion or use of response styles as a shortcut method may also influence likelihood of extreme response style. Regarding the measurement of extreme style, there seems to be a latent construct of response extremity that generalizes across differing item content, and although the measurement of extreme response style was not improved through the use of Mixed Rasch methods, it is possible that given a large enough sample size, ERS can be accurately identified both through simple proportional and latent-class based techniques.

An accumulation of evidence in this study thus indicates that although extreme response style may be influenced by situational factors, a substantial amount of the
variance in extreme responding is due to stable personality traits. This conclusion is buttressed by support from moderate relations between extreme response style and two personality variables, evidence that a latent construct of response extremity occurs across measures irrespective of specific item content, and the lack of fatigue as an explanation for extreme response style. The implications of these results have bearing in any work situation that requires the rating of a stimulus or individual, as it appears that personality factors may play a substantial role in whether or not an extreme response is chosen over a non-extreme response, irrespective of the actual content of the stimulus. In situations where cautious judgments are preferable, or where the consequences of an inaccurate extreme judgment may be severe (such as in clinical or medical decisions) it may be necessary to ensure that the rater does not possess any personality characteristics that may be associated with an extreme response style.

In selecting the potential research questions to explore for this study, I have tried to be as thorough as possible in exploring two overarching goals: determining the best way to measure extreme response style, and exploring relevant personality correlates that may predict extreme response style. Understanding response styles such as ERS is not only important from a psychometric or practical standpoint, but also interesting as a theoretical investigation of a dispositional style that may influence behavior in substantive areas of work (Guion, 1998). Casting extreme response style as a behavioral manifestation of intolerance of ambiguity and decisiveness also serves as a form of validation for these personality constructs, based on the hypothesized links between dispositional traits and behavioral actions. Future research should improve upon the limitations in this study inherent in its correlational design, and seek to establish a more causal link between these
personality traits and extreme response style in a laboratory setting. Given that there are countless hypotheses or relations that could be tested based on this data-set, I have tried to select hypotheses that I find interesting and to pursue research questions that present some relevance to the practical applications of studying extreme responding in an organizational context. Within this wealth of data, hopefully several interesting or surprising conclusions have emerged that illuminate and contribute to the intolerance of ambiguity, extreme response style, organizational survey, and general psychometric literatures.
References


Brengelmann, J. (1960a). Extreme response set, drive level and abnormality in


Budner, S. (1962). Intolerance of ambiguity as a personality variable. *Journal of
Personality*, 30, 29-59.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative
space: A critical review, with emphasis on the separability of positive and negative

investigation. *Journal of Social Behavior and Personality*, 15, 137-152.

of use of extreme positive and negative ratings. *Journal of Social Psychology*, 89,
281-293.

Measurement*, 6, 475-494.

Cronbach, L. (1950). Further evidence on response sets and test design. *Educational and
Psychological Measurement*, 10, 3-31.


Gough, H. & Sanford, R. (1952). Rigidity as a psychological variable. (Unpublished manuscript, Institute of Personality Assessment and Research, University of California, Berkeley.)


Table 1 *List of All Personality Measures in Questionnaire Battery, by Section*

**Section 1:**
- Need for Closure*
- Personal Fear of Invalidity*
- Need for Structure
- Need for Cognition
- Social Dominance Orientation
- Openness to Experience
- Agreeableness
- Conscientiousness

**Section 2:**
- Conservatism
- Cognitive Complexity
- Reactions to Ambiguity*
- Simplistic Thinking
- Decisiveness (Seizing)*
- Confidence
- Resistance to Change
- Close-mindedness
- Fear of Death
- Extremeness

**Section 3:**
- Greenleaf (1992) Measure of Extreme Response Style*
- Rosenberg (1965) Self-Esteem Scale
- Resume Ratings*

Note: Measures relevant to the hypothesized relations are marked with *
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
<th>7.</th>
<th>8.</th>
<th>9.</th>
<th>10.</th>
<th>11.</th>
<th>12.</th>
<th>13.</th>
<th>14.</th>
<th>15.</th>
<th>16.</th>
<th>17.</th>
<th>18.</th>
<th>19.</th>
<th>20.</th>
<th>21.</th>
<th>22.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Gender</td>
<td>1.61</td>
<td>0.49</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Ethnicity</td>
<td>1.14</td>
<td>0.35</td>
<td>-0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Time</td>
<td>2783</td>
<td>1325</td>
<td>-0.03</td>
<td>0.17</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. SAT</td>
<td>1198</td>
<td>125</td>
<td>-0.22</td>
<td>-0.05</td>
<td>0.05</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Complexity</td>
<td>3.69</td>
<td>0.76</td>
<td>-0.10</td>
<td>-0.02</td>
<td>-0.15</td>
<td>-0.14</td>
<td>(0.83)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Contradiction</td>
<td>3.50</td>
<td>0.84</td>
<td>-0.05</td>
<td>-0.16</td>
<td>-0.09</td>
<td>-0.14</td>
<td>0.73</td>
<td>(0.69)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Vagueness</td>
<td>4.20</td>
<td>0.67</td>
<td>-0.01</td>
<td>-0.11</td>
<td>-0.13</td>
<td>-0.08</td>
<td>0.70</td>
<td>0.70</td>
<td>(0.79)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Ambiguity Mean</td>
<td>3.90</td>
<td>0.65</td>
<td>-0.06</td>
<td>-0.08</td>
<td>-0.14</td>
<td>-0.11</td>
<td>0.91</td>
<td>0.84</td>
<td>0.90</td>
<td>(0.91)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Peer Complexity</td>
<td>3.87</td>
<td>0.78</td>
<td>-0.08</td>
<td>-0.04</td>
<td>-0.16</td>
<td>-0.15</td>
<td>0.46</td>
<td>0.41</td>
<td>0.44</td>
<td>0.49</td>
<td>(0.87)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Peer Contradiction</td>
<td>3.76</td>
<td>0.78</td>
<td>-0.13</td>
<td>-0.16</td>
<td>-0.14</td>
<td>0.13</td>
<td>0.38</td>
<td>0.37</td>
<td>0.46</td>
<td>0.46</td>
<td>0.65</td>
<td>(0.67)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Peer Vagueness</td>
<td>4.28</td>
<td>0.75</td>
<td>0.03</td>
<td>-0.13</td>
<td>-0.05</td>
<td>-0.02</td>
<td>0.34</td>
<td>0.37</td>
<td>0.38</td>
<td>0.40</td>
<td>0.70</td>
<td>0.66</td>
<td>(0.87)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Peer Ambiguity Mean</td>
<td>4.05</td>
<td>0.67</td>
<td>-0.06</td>
<td>-0.10</td>
<td>-0.12</td>
<td>-0.02</td>
<td>0.46</td>
<td>0.45</td>
<td>0.49</td>
<td>0.52</td>
<td>0.89</td>
<td>0.82</td>
<td>0.91</td>
<td>(0.93)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Decisiveness</td>
<td>3.36</td>
<td>0.78</td>
<td>-0.24</td>
<td>0.02</td>
<td>-0.15</td>
<td>0.07</td>
<td>0.22</td>
<td>0.31</td>
<td>0.21</td>
<td>0.26</td>
<td>0.14</td>
<td>0.11</td>
<td>0.00</td>
<td>0.10</td>
<td>(0.80)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Peer Decisiveness</td>
<td>3.59</td>
<td>0.73</td>
<td>-0.17</td>
<td>0.00</td>
<td>-0.08</td>
<td>0.08</td>
<td>0.20</td>
<td>0.14</td>
<td>0.10</td>
<td>0.18</td>
<td>0.12</td>
<td>0.10</td>
<td>0.11</td>
<td>0.15</td>
<td>0.55</td>
<td>(0.82)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. ERS Greenleaf</td>
<td>0.26</td>
<td>0.19</td>
<td>-0.02</td>
<td>0.10</td>
<td>-0.04</td>
<td>0.03</td>
<td>0.16</td>
<td>0.06</td>
<td>0.09</td>
<td>0.14</td>
<td>0.28</td>
<td>0.14</td>
<td>0.31</td>
<td>0.30</td>
<td>0.15</td>
<td>0.25</td>
<td>(0.71)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. ERS Alternate</td>
<td>0.20</td>
<td>0.15</td>
<td>-0.19</td>
<td>0.09</td>
<td>-0.04</td>
<td>0.11</td>
<td>0.18</td>
<td>0.13</td>
<td>0.26</td>
<td>0.23</td>
<td>0.32</td>
<td>0.19</td>
<td>0.24</td>
<td>0.28</td>
<td>0.29</td>
<td>0.22</td>
<td>0.45</td>
<td>(0.67)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. ERS Overall</td>
<td>0.20</td>
<td>0.12</td>
<td>0.03</td>
<td>0.05</td>
<td>-0.02</td>
<td>0.19</td>
<td>0.06</td>
<td>0.09</td>
<td>0.26</td>
<td>0.18</td>
<td>0.25</td>
<td>0.18</td>
<td>0.27</td>
<td>0.29</td>
<td>0.17</td>
<td>0.25</td>
<td>0.50</td>
<td>0.70</td>
<td>(0.97)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. ERS Resume Rating</td>
<td>0.21</td>
<td>0.27</td>
<td>-0.04</td>
<td>-0.14</td>
<td>-0.04</td>
<td>-0.03</td>
<td>0.13</td>
<td>0.22</td>
<td>0.29</td>
<td>0.24</td>
<td>0.08</td>
<td>0.15</td>
<td>0.09</td>
<td>0.13</td>
<td>0.02</td>
<td>0.04</td>
<td>0.17</td>
<td>0.12</td>
<td>0.38</td>
<td>(0.91)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. ERS Unit-weighted</td>
<td>0.22</td>
<td>0.13</td>
<td>-0.07</td>
<td>0.10</td>
<td>-0.04</td>
<td>0.12</td>
<td>0.17</td>
<td>0.11</td>
<td>0.23</td>
<td>0.22</td>
<td>0.35</td>
<td>0.35</td>
<td>0.24</td>
<td>0.29</td>
<td>0.83</td>
<td>0.83</td>
<td>0.84</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. ERS Mixed Rasch</td>
<td>0.28</td>
<td>0.39</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.08</td>
<td>0.22</td>
<td>-0.10</td>
<td>-0.08</td>
<td>0.01</td>
<td>-0.05</td>
<td>0.02</td>
<td>0.06</td>
<td>0.08</td>
<td>0.07</td>
<td>0.12</td>
<td>0.05</td>
<td>0.18</td>
<td>0.18</td>
<td>0.42</td>
<td>0.23</td>
<td>0.23</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. Conscientiousness</td>
<td>0.24</td>
<td>0.39</td>
<td>-0.02</td>
<td>0.06</td>
<td>0.00</td>
<td>-0.10</td>
<td>-0.04</td>
<td>0.06</td>
<td>-0.02</td>
<td>-0.03</td>
<td>0.07</td>
<td>-0.11</td>
<td>0.04</td>
<td>0.01</td>
<td>0.21</td>
<td>0.28</td>
<td>0.22</td>
<td>0.36</td>
<td>0.51</td>
<td>0.16</td>
<td>0.41</td>
<td>0.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. ERS Mixed Rasch</td>
<td>0.26</td>
<td>0.40</td>
<td>-0.08</td>
<td>-0.05</td>
<td>0.14</td>
<td>0.03</td>
<td>-0.08</td>
<td>-0.05</td>
<td>0.00</td>
<td>-0.04</td>
<td>-0.01</td>
<td>-0.01</td>
<td>0.04</td>
<td>0.03</td>
<td>0.01</td>
<td>0.04</td>
<td>0.24</td>
<td>0.25</td>
<td>0.43</td>
<td>0.08</td>
<td>0.35</td>
<td>0.25</td>
<td>0.31</td>
<td></td>
</tr>
</tbody>
</table>

Note: Reliabilities are listed on the main diagonal of the matrix. N-values range from 108-116 participants.
* Correlation is significant at the 0.05 level (2-tailed).
<table>
<thead>
<tr>
<th>Factors</th>
<th>TAMU 1</th>
<th>TAMU 2</th>
<th>TAMU 3</th>
<th>Rice 1</th>
<th>Rice 2</th>
<th>Rice 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thinking about complex brain teasers is a waste of time.</td>
<td>.53</td>
<td>.17</td>
<td>.33</td>
<td>.20</td>
<td>.62</td>
<td>.10</td>
</tr>
<tr>
<td>I find it annoying when someone says something that is overly complex.</td>
<td>.61</td>
<td>.21</td>
<td>.37</td>
<td>.20</td>
<td>.63</td>
<td>.20</td>
</tr>
<tr>
<td>I have no use for people who have 'nuanced' opinions.</td>
<td>.66</td>
<td>.25</td>
<td>.28</td>
<td>.35</td>
<td>.74</td>
<td>.04</td>
</tr>
<tr>
<td>I don't like things that have more than one interpretation.</td>
<td>.76</td>
<td>.35</td>
<td>.46</td>
<td>.67</td>
<td>.52</td>
<td>.14</td>
</tr>
<tr>
<td>I like things that are plain and easy to understand.</td>
<td>.41</td>
<td>.43</td>
<td>.73</td>
<td>.50</td>
<td>.40</td>
<td>.43</td>
</tr>
<tr>
<td>The simple life is the good life.</td>
<td>.27</td>
<td>.29</td>
<td>.68</td>
<td>.48</td>
<td>-.01</td>
<td>.30</td>
</tr>
<tr>
<td>I prefer simple problems over complex problems.</td>
<td>.42</td>
<td>.34</td>
<td>.83</td>
<td>.62</td>
<td>.43</td>
<td>.38</td>
</tr>
<tr>
<td>I like it when issues are black and white rather than 'shades of grey.'</td>
<td>.50</td>
<td>.46</td>
<td>.81</td>
<td>.76</td>
<td>.31</td>
<td>.21</td>
</tr>
<tr>
<td>I like it when there are no ifs, ands, or buts about it.</td>
<td>.54</td>
<td>.43</td>
<td>.74</td>
<td>.70</td>
<td>.42</td>
<td>.31</td>
</tr>
<tr>
<td>Poems with contradictions are annoying.</td>
<td>.69</td>
<td>.43</td>
<td>.37</td>
<td>.43</td>
<td>.68</td>
<td>.18</td>
</tr>
<tr>
<td>I dislike things that have both positive and negative qualities about them.</td>
<td>.64</td>
<td>.30</td>
<td>.39</td>
<td>.43</td>
<td>.62</td>
<td>.12</td>
</tr>
<tr>
<td>Groups which have too much difference of opinion among their own members don't appeal to me.</td>
<td>.56</td>
<td>.41</td>
<td>.33</td>
<td>.47</td>
<td>.55</td>
<td>.20</td>
</tr>
<tr>
<td>I like stories that have consistent characters.</td>
<td>.44</td>
<td>.54</td>
<td>.39</td>
<td>.59</td>
<td>.25</td>
<td>.36</td>
</tr>
<tr>
<td>The most important quality in a person is consistency of opinions.</td>
<td>.45</td>
<td>.26</td>
<td>.22</td>
<td>.41</td>
<td>.07</td>
<td>.25</td>
</tr>
<tr>
<td>Vague and impressionistic pictures really have little appeal for me.</td>
<td>.69</td>
<td>.28</td>
<td>.24</td>
<td>.62</td>
<td>.33</td>
<td>-.04</td>
</tr>
<tr>
<td>I don't like to work on a problem unless there's a possibility of coming out with a clear cut answer.</td>
<td>.62</td>
<td>.58</td>
<td>.44</td>
<td>.71</td>
<td>.47</td>
<td>.21</td>
</tr>
<tr>
<td>I hate it when you can't solve a problem right away.</td>
<td>.55</td>
<td>.51</td>
<td>.40</td>
<td>.58</td>
<td>.34</td>
<td>.50</td>
</tr>
<tr>
<td>I hate it when a TV series ends its season with a cliffhanger.</td>
<td>.33</td>
<td>.52</td>
<td>.24</td>
<td>.20</td>
<td>.49</td>
<td>.48</td>
</tr>
<tr>
<td>A group meeting functions best with a definite agenda.</td>
<td>.17</td>
<td>.65</td>
<td>.30</td>
<td>-.04</td>
<td>.10</td>
<td>.50</td>
</tr>
<tr>
<td>The best part of working a jigsaw puzzle is putting in that last piece.</td>
<td>.28</td>
<td>.56</td>
<td>.19</td>
<td>.04</td>
<td>.07</td>
<td>.53</td>
</tr>
<tr>
<td>I like movies with clear cut endings.</td>
<td>.33</td>
<td>.66</td>
<td>.31</td>
<td>.33</td>
<td>.23</td>
<td>.58</td>
</tr>
<tr>
<td>A good job is one where what is to be done is always clear.</td>
<td>.39</td>
<td>.69</td>
<td>.48</td>
<td>.43</td>
<td>.26</td>
<td>.58</td>
</tr>
<tr>
<td>I usually get a strong sense of relief when I finally get the answer.</td>
<td>.06</td>
<td>.64</td>
<td>.38</td>
<td>.32</td>
<td>.10</td>
<td>.62</td>
</tr>
<tr>
<td>I'd rather find an answer, any answer, compared to uncertainty.</td>
<td>.37</td>
<td>.60</td>
<td>.23</td>
<td>.26</td>
<td>.04</td>
<td>.70</td>
</tr>
<tr>
<td>I dislike questions which could be answered in many different ways.</td>
<td>.51</td>
<td>.44</td>
<td>.35</td>
<td>.59</td>
<td>.29</td>
<td>-.06</td>
</tr>
</tbody>
</table>
Table 4 Differing Factor Models of the Reactions to Ambiguity Scale

<table>
<thead>
<tr>
<th>Number of Factors</th>
<th>Number of Items</th>
<th>Alpha</th>
<th>CFI</th>
<th>SRMR</th>
<th>Chi-square</th>
<th>df</th>
<th>Difference</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25</td>
<td>.91</td>
<td>.75</td>
<td>.071</td>
<td>1202.69</td>
<td>275</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 (Complete)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complexity</td>
<td>10</td>
<td>.83</td>
<td>.80</td>
<td>.066</td>
<td>1005.01</td>
<td>272</td>
<td>p&lt;.01</td>
<td></td>
</tr>
<tr>
<td>Contradiction</td>
<td>5</td>
<td>.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vagueness</td>
<td>10</td>
<td>.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 (Reduced)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complexity</td>
<td>4</td>
<td>.81</td>
<td>.93</td>
<td>.045</td>
<td>179.31</td>
<td>51</td>
<td>p&lt;.01</td>
<td></td>
</tr>
<tr>
<td>Contradiction</td>
<td>4</td>
<td>.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vagueness</td>
<td>4</td>
<td>.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5 Consistent Akaike's Information Criterion Values Determining Number of Classes

<table>
<thead>
<tr>
<th>Number of Classes</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conscientiousness</td>
<td>12367.60</td>
<td>12187.45</td>
<td>12247.76</td>
</tr>
<tr>
<td>Need For Structure</td>
<td>15074.39</td>
<td>14891.06</td>
<td>15048.82</td>
</tr>
<tr>
<td>Need For Cognition</td>
<td>21973.53</td>
<td>21689.80</td>
<td>21887.39</td>
</tr>
</tbody>
</table>

Note. Values in italics indicate the best-fitting solution.
Figure 1. The Moderating Effect of Decisiveness on the Relationship between Time and Extreme Response Style (Predicted Effect)
Figure 2. The interactive effect of decisiveness and intolerance of ambiguity on extreme responding (Self-Report)

Figure 3. The interactive effect of decisiveness and intolerance of ambiguity on extreme responding (Peer-Report)
Figure 4. The interactive effect of time spent on survey and decisiveness (seizing) on extreme responding (Self-Report)

Figure 5. The interactive effect of time spent on survey and decisiveness (seizing) on extreme responding (Peer-Report)
Figure 6. Threshold Parameters for Non-Extreme Class of Conscientiousness Scale

Figure 7. Threshold Parameters for Extreme Class of Conscientiousness Scale
Figure 8. Threshold Parameters for Non-Extreme Class of Need for Structure Scale

Figure 9. Threshold Parameters for Extreme Class of Need for Structure Scale
Figure 10. Threshold Parameters for Non-Extreme Class of Need for Cognition Scale

Figure 11. Threshold Parameters for Extreme Class of Need for Cognition Scale
Appendix

Instructions for Participants:

In this study, you will be asked to complete a series of questionnaires which should take no more than an hour of your time. Please answer all questions as honestly as you can. There are no right or wrong answers so do not spend too much time deciding on an answer. The first thing that comes to mind is probably the best response. There is no time limit, but please work as quickly as possible. After completing the questionnaire, please indicate whether you will also be having a peer who knows you well complete the study for one hour of additional credit.

Instructions for Peers

In this study, you will be asked to complete a series of questionnaires which should take no more than an hour of your time. Please answer all questions AS IF YOU WERE YOUR FRIEND/THE PARTICIPANT OF THE EXPERIMENT. Do NOT answer the questions with your own personal responses. If you feel you do not know how your friend might respond to a particular question, or you feel that you don’t have enough information to answer a question, please take a guess as to how you feel your friend might respond. There is no time limit, but please work as quickly as possible.

Both participants will complete the same survey, except the primary participant will be asked to rate the resume. The peer will not be asked to rate the resume.

PREVIOUSLY VALIDATED INSTRUMENTS:

Need for Closure Scale (NFCS: Webster & Kruglanski, 1994)

1 = strongly disagree; 2 = moderately disagree; 3 = slightly disagree; 4 = slightly agree; 5 = moderately agree; 6 = strongly agree

1. I think that having clear rules and order at work is essential for success.
2. Even after I’ve made up my mind about something, I am always eager to consider a different opinion.
3. I don’t like situations that are uncertain.
4. I dislike questions which could be answered in many different ways.
5. I like to have friends who are unpredictable.
6. I find that a well-ordered life with regular hours suits my temperament.
7. When dining out, I like to go to places where I have been before so that I know what to expect.
8. I feel uncomfortable when I don’t understand why an event occurred in my life.
9. I feel irritated when one person disagrees with what everyone else in a group believes.
10. I hate to change my plans at the last minute.
11. I don’t like to go into a situation without knowing what I can expect from it.
12. When I go shopping, I have difficulty deciding exactly what it is that I want.
13. When faced with a problem I usually see the one best solution very quickly.
14. When I am confused about an important issue, I feel very upset.
15. I tend to put off making important decisions until the last possible moment.
16. I usually make important decisions quickly and confidently.
17. I would describe myself as indecisive.
18. I think it is fun to change my plans at the last minute.
19. I enjoy the uncertainty of going into a new situation without knowing what might happen.
20. My personal space is usually messy and disorganized.
21. In most social conflicts, I can easily see which side is right and which is wrong.
22. I tend to struggle with most decisions.
23. I believe that orderliness and organization are among the most important characteristics of a good student.
24. When considering most conflict situations, I can usually see how both sides could be right.
25. I don't like to be with people who are capable of unexpected actions.
26. I prefer to socialize with familiar friends because I know what to expect from them.
27. I think that I would learn best in a class that lacks clearly stated objectives and requirements.
28. When thinking about a problem, I consider as many different opinions on the issue as possible.
29. I like to know what people are thinking all the time.
30. I dislike it when a person's statement could mean many different things.
31. It's annoying to listen to someone who cannot seem to make up his or her mind.
32. I find that establishing a consistent routine enables me to enjoy life more.
33. I enjoy having a clear and structured mode of life.
34. I prefer interacting with people whose opinions are very different from my own.
35. I like to have a place for everything and everything in its place.
36. I feel uncomfortable when someone's meaning or intention is unclear to me.
37. When trying to solve a problem I often see so many possible options that it's confusing.
38. I always see so many possible solutions to problems I face.
39. I'd rather know bad news than stay in a state of uncertainty.
40. I do not usually consult many different opinions before forming my own view.
41. I dislike unpredictable situations.
42. I dislike the routine aspects of my work (studies).

**NFCS Discomfort with Ambiguity Facet:**

I don't like situations that are uncertain.
I feel uncomfortable when I don't understand why an event occurred in my life.
When I am confused about an important issue, I feel very upset.
In most social conflicts, I can easily see which side is right and which is wrong.
I like to know what people are thinking all the time.
I dislike it when a person's statement could mean many different things.
It's annoying to listen to someone who cannot seem to make up his or her mind. I'd rather know bad news than stay in a state of uncertainty.

NFCS Decisiveness Facet:

When I go shopping, I have difficulty deciding exactly what it is that I want. *(reverse scored)*
When faced with a problem I usually see the one best solution very quickly. I tend to put off making important decisions until the last possible moment. *(reverse scored)*
I usually make important decisions quickly and confidently. I would describe myself as indecisive. *(reverse scored)*
I tend to struggle with most decisions. *(reverse scored)*
When trying to solve a problem I often see so many possible options that it's confusing. *(reverse scored)*

Personal Fear of Invalidity Scale (PFI: Thomson et al, 1989)

1. I may struggle with a few decisions, but not very often. *(reverse scored)*
2. I never put off making important decisions. *(reverse scored)*
3. Sometimes I become impatient over my indecisiveness.
4. Sometimes I see so many options to a situation that it is really confusing.
5. I can be reluctant to commit myself to something because of the possibility that I might be wrong.
6. I tend to struggle with most decisions.
7. Even after making an important decision, I continue to think about the pros and cons to make sure that I am not wrong.
8. Regardless of whether others see an event as positive or negative, I don't mind committing myself to it. *(reverse scored)*
9. I prefer situations where I don't have to decide immediately.
10. I rarely doubt that the course of action I have selected will be correct. *(reverse scored)*
11. I tend to continue to evaluate recently made decisions.
12. I wish I didn't worry so much about making errors.
13. Decisions rarely weigh heavily on my shoulders. *(reverse scored)*
14. I find myself reluctant to commit to new ideas but find little comfort in remaining with the tried and true.
Extreme Response Style (Greenleaf, 1992)

1- definitely disagree;
2- moderately disagree;
3- slightly disagree
4- slightly agree
5- moderately agree
6- definitely agree

1. When I see a full ashtray or wastebasket, I want to empty it immediately.
2. I am a homebody.
3. Television is my primary form of entertainment.
4. No matter how fast our income goes up, we never seem to get ahead.
5. I try to avoid foods that are high in cholesterol.
6. Advertising insults my intelligence.
7. Investing in the stock market is too risky for most families.
8. Everyone should use a mouthwash to help control bad breath.
9. TV commercials place too much emphasis on sex.
10. A college education is very important for success in today's world.
11. My days seem to follow a definite routine-eating meals at the same time each day, etc.
12. I like to visit places that are totally different from my home.
13. I work very hard most of the time.
14. I like to feel attractive to members of the opposite sex.
15. I will probably have more money to spend next year than I have now.
16. I eat more than I should.

NEWLY CREATED SCALES:

Items marked with an asterisk (*) also appear in the previously validated measures, and thus, will not be repeated.

Reactions to Ambiguity

1 (Strongly Disagree), 2 (Disagree), 3 (Slightly Disagree), 4 (Slightly Agree), 5 (Agree), 6 (Strongly Agree)

Global reactions to ambiguity:
1. I dislike things that are ambiguous.
2. I like it when things are clear cut.
3. I don't like people that are wishy-washy.

Complexity:
Discomfort with complexity
4. Thinking about complex brainteasers is a waste of time.
5. I dislike questions which could be answered correctly in many different ways. (NFCS)
6. I find it annoying when someone says something that is overly complex.
7. I have no use for people who have "nuanced" opinions.
8. I don't like things that have more than one interpretation.

Preference for simplicity
9. I like things that are plain and easy to understand.
10. The simple life is the good life.
11. I prefer simple problems over complex problems.
12. I like it when issues are black and white rather than "shades of grey."
13. I like it when there are no ifs, ands, or buts about it.

Contradiction:
Discomfort with contradiction
14. It really bothers me to choose between two equally good alternatives.
15. Poems with contradictions are annoying.
16. I dislike things that have both positive and negative qualities about them.
17. Groups which have too much difference of opinion among their own members don't appeal to me.

Preference for consistency
18. I like stories that have consistent characters.
19. The most important quality in a person is consistency of opinions.

Vagueness
Discomfort with vagueness
20. Vague and impressionistic pictures really have little appeal for me.
21. I don't like to work on a problem unless there's a possibility of coming out with a clear cut answer.
22. It's annoying to listen to someone who cannot seem to make up his or her mind.* (NFCS)
23. I hate it when you can't solve a problem right away.
24. I hate it when a TV series ends its season with a cliffhanger.

Preference for completeness
25. A group meeting functions best with a definite agenda.
26. The best part of working a jigsaw puzzle is putting in that last piece.
27. I'd rather hear bad news than no news at all.
28. I like movies with clear cut endings.
29. A good job is one where what is to be done is always clear.
30. I usually get a strong sense of relief when I finally get the answer.
31. I'd rather find an answer, any answer, compared to uncertainty.
32. When I start reading a novel, I usually can't put it down until finish it.
Decisiveness

1. I may struggle with a few decisions, but not very often.* (PFI)
2. I never put off making important decisions.* (PFI)
3. I can be reluctant to commit myself to something because of the possibility that I might be wrong.
4. The possibility that I might be wrong rarely stops me from making a decision.
5. Regardless of whether others see an event as positive or negative, I don't mind committing myself to it.* (PFI)
6. I tend to make my decisions immediately.
7. When faced with a problem I usually see the one best solution very quickly.*
8. I usually make important decisions quickly.
9. I would describe myself as decisive.
10. He who hesitates is lost.
Resume Rating

For the next activity, please pretend that you are the head of Human Resources for a large organization. You are trying to fill a position for a new HR manager, and have received a variety of resumes. One such resume appears below. Please review the resume and evaluate it using the questions that immediately follow. Please try to weight the experience, education, and skills of the applicant evenly when making your evaluation.

Resume Sample:

John Brown

Objective
To obtain management position in human resources where my years of experience and skills will be optimally employed.


Experience

Human Resources Manager

- Oversaw the interviewing, hiring and training of new employees.
- Supervised Human Resources Specialists and clerks.
- Handled employee grievances.
- Formulated the budget for the Human Resources Department.

1996-1999 Clemson Group Chapel Hill, NC

Human Resources Specialist

- Led training sessions about team building, sexual harassment, diversity, and conflict resolution.
- Interviewed prospective employees.
- Acted as a liaison for new hires.

1994-1996 Clemson Group Chapel Hill, NC

Benefits Specialist

- Generated and audited report of benefit plan members
- Processed and maintained compliance for leave policies
- Researched and recommended improvements to benefits program

1993-1994 Clemson Group Chapel Hill, NC

Benefits Assistant

- Compiled and distribute benefits packages.
- Maintained benefit files
- Maintained employee changes in benefits.

1989–1993 Duke University Durham,
NC

**Student Library Assistant**
- Assisted students with questions regarding the library.
- Checked books in and out.
- Reshelved books.

Summer 1992 Clemson Group Chapel Hill, NC

**Summer Intern**
- Worked with benefits administrator.
- Helped file and process claims.
- Provided general office support.

**Education**
1996-1998 UNC Kenan-Flagler Business School Chapel Hill, NC
- MBA
- Cumulative GPA: 3.5

1989-1993 Duke University Durham, NC
- B.A., Business Administration and Psychology
- Cumulative GPA: 2.95

**Skills**
Excellent written and oral communication skills
Proficient with the Microsoft office suite
Trained in conflict resolution techniques

**Reference**
Available upon request

John Brown should be hired for the job.

John Brown fits the job description well.

John Brown is well qualified for this job.

John Brown will succeed in the business field.

John Brown is likely to be called back for an interview.

John Brown has a good educational background for this job.

John Brown has good work experience for this job.

John Brown is unlikely to get a job in this field.

How would you rate this resume in terms of aesthetic quality/attractiveness?

How would you rate the quality of this candidate as a whole?
Means and Standard Deviations: TAMU Summer Sample Overall

<table>
<thead>
<tr>
<th>Descriptive Statistics</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender:</td>
<td>1.60</td>
<td>.490</td>
<td>232</td>
</tr>
<tr>
<td>ETHNIC2</td>
<td>.26</td>
<td>.438</td>
<td>233</td>
</tr>
<tr>
<td>SAT</td>
<td>1135.26</td>
<td>137.500</td>
<td>178</td>
</tr>
<tr>
<td>TIME</td>
<td>2574.8533</td>
<td>1285.78555</td>
<td>225</td>
</tr>
<tr>
<td>COMPLEX</td>
<td>3.6010</td>
<td>.75822</td>
<td>233</td>
</tr>
<tr>
<td>CONTRA</td>
<td>3.5028</td>
<td>.77950</td>
<td>232</td>
</tr>
<tr>
<td>VAGUE</td>
<td>4.1765</td>
<td>.70516</td>
<td>232</td>
</tr>
<tr>
<td>SEIZE</td>
<td>3.4570</td>
<td>.73732</td>
<td>231</td>
</tr>
<tr>
<td>GREEN</td>
<td>.2555</td>
<td>.18047</td>
<td>231</td>
</tr>
<tr>
<td>ALTGREEN</td>
<td>.2022</td>
<td>.15661</td>
<td>231</td>
</tr>
<tr>
<td>ERSMINUS</td>
<td>.1724</td>
<td>.12052</td>
<td>231</td>
</tr>
<tr>
<td>RATEMEAN</td>
<td>.1709</td>
<td>.27101</td>
<td>231</td>
</tr>
</tbody>
</table>

Means and Standard Deviations: TAMU Summer Sample Peers Only

<table>
<thead>
<tr>
<th>Descriptive Statistics</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender:</td>
<td>1.58</td>
<td>.497</td>
<td>62</td>
</tr>
<tr>
<td>ETHNIC2</td>
<td>.27</td>
<td>.450</td>
<td>62</td>
</tr>
<tr>
<td>SAT</td>
<td>1126.96</td>
<td>129.942</td>
<td>46</td>
</tr>
<tr>
<td>TIME</td>
<td>2852.4237</td>
<td>1362.76523</td>
<td>59</td>
</tr>
<tr>
<td>COMPLEX</td>
<td>3.7692</td>
<td>.71110</td>
<td>62</td>
</tr>
<tr>
<td>CONTRA</td>
<td>3.5742</td>
<td>.72110</td>
<td>62</td>
</tr>
<tr>
<td>VAGUE</td>
<td>4.2147</td>
<td>.67563</td>
<td>62</td>
</tr>
<tr>
<td>SEIZE</td>
<td>3.5205</td>
<td>.67669</td>
<td>61</td>
</tr>
<tr>
<td>PCOMPLEX</td>
<td>3.6509</td>
<td>.62919</td>
<td>62</td>
</tr>
<tr>
<td>PCONTRA</td>
<td>3.7177</td>
<td>.77068</td>
<td>62</td>
</tr>
<tr>
<td>PVAGUE</td>
<td>4.1308</td>
<td>.61656</td>
<td>62</td>
</tr>
<tr>
<td>PSEIZE</td>
<td>3.6246</td>
<td>.73101</td>
<td>62</td>
</tr>
<tr>
<td>GREEN</td>
<td>.2691</td>
<td>.18516</td>
<td>62</td>
</tr>
<tr>
<td>ALTGREEN</td>
<td>.2033</td>
<td>.16087</td>
<td>62</td>
</tr>
<tr>
<td>ERSMINUS</td>
<td>.1656</td>
<td>.11181</td>
<td>62</td>
</tr>
<tr>
<td>RATEMEAN</td>
<td>.1645</td>
<td>.26058</td>
<td>62</td>
</tr>
</tbody>
</table>
Means and Standard Deviations: Rice Sample

<table>
<thead>
<tr>
<th>Descriptive Statistics</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender:</td>
<td>1.60</td>
<td>.492</td>
<td>127</td>
</tr>
<tr>
<td>ETHNIC2</td>
<td>.4557</td>
<td>.50009</td>
<td>127</td>
</tr>
<tr>
<td>SAT</td>
<td>1371.4783</td>
<td>133.39194</td>
<td>115</td>
</tr>
<tr>
<td>TIME</td>
<td>2633.9426</td>
<td>1551.19706</td>
<td>122</td>
</tr>
<tr>
<td>COMPLEX</td>
<td>3.2738</td>
<td>.66496</td>
<td>127</td>
</tr>
<tr>
<td>CONTRA</td>
<td>3.0909</td>
<td>.78693</td>
<td>126</td>
</tr>
<tr>
<td>VAGUE</td>
<td>3.8305</td>
<td>.60087</td>
<td>126</td>
</tr>
<tr>
<td>SEIZE</td>
<td>3.1261</td>
<td>.79304</td>
<td>127</td>
</tr>
<tr>
<td>GREEN</td>
<td>.2170</td>
<td>.19010</td>
<td>124</td>
</tr>
<tr>
<td>ALTGREEN</td>
<td>.2197</td>
<td>.14381</td>
<td>127</td>
</tr>
<tr>
<td>ERSMINUS</td>
<td>.2111</td>
<td>.12747</td>
<td>127</td>
</tr>
<tr>
<td>RATEMEAN</td>
<td>.0921</td>
<td>.20070</td>
<td>125</td>
</tr>
<tr>
<td>Gender</td>
<td>ETHNIC</td>
<td>SAT</td>
<td>TIME</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>-----</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENDER</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>-1.74</td>
<td>1</td>
<td>-1.47</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>0.08</td>
<td>0.23</td>
<td>0.18</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>-0.84</td>
<td>-0.56</td>
<td>-0.62</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>-0.56</td>
<td>-0.62</td>
<td>-0.71</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>-0.62</td>
<td>-0.71</td>
<td>-0.55</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>-0.71</td>
<td>-0.55</td>
<td>-0.62</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>-0.55</td>
<td>-0.62</td>
<td>-0.71</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>-0.62</td>
<td>-0.71</td>
<td>-0.55</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>-0.71</td>
<td>-0.55</td>
<td>-0.62</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>-0.55</td>
<td>-0.62</td>
<td>-0.71</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>-0.62</td>
<td>-0.71</td>
<td>-0.55</td>
</tr>
<tr>
<td>Sig (2-tailed)</td>
<td>-0.71</td>
<td>-0.55</td>
<td>-0.62</td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.01 level (2-tailed).**Correlation is significant at the 0.05 level (2-tailed).
## Correlations: TAMU Summer Sample Peers Only

<table>
<thead>
<tr>
<th>Gender</th>
<th>Pearson Correlation Sig. (2-tailed)</th>
<th>SAT</th>
<th>TIME</th>
<th>COMPLEX</th>
<th>CONTRA</th>
<th>VAGUE</th>
<th>SIZE</th>
<th>PCOMPLEX</th>
<th>PCONTRA</th>
<th>PVAGUE</th>
<th>PSEIZE</th>
<th>GREEN</th>
<th>ALTGREEN</th>
<th>ERSMINUS</th>
<th>RATEMEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.19*</td>
<td>-0.196*</td>
<td>0.021</td>
<td>1</td>
<td>-0.13**</td>
<td>0.037**</td>
<td>0.23</td>
<td>-0.12**</td>
<td>0.1**</td>
<td>0.1**</td>
<td>-0.07*</td>
<td>0.000</td>
<td>-0.04**</td>
<td>-0.04**</td>
<td>0.000</td>
</tr>
<tr>
<td>Female</td>
<td>-0.13</td>
<td>-0.087</td>
<td>0.046</td>
<td>0.072</td>
<td>-0.12**</td>
<td>0.016</td>
<td>0.1**</td>
<td>-0.12**</td>
<td>0.1**</td>
<td>0.1**</td>
<td>-0.07*</td>
<td>0.000</td>
<td>-0.04**</td>
<td>-0.04**</td>
<td>0.000</td>
</tr>
<tr>
<td>ETHNIC2</td>
<td>-0.25</td>
<td>-0.201</td>
<td>-0.059</td>
<td>0.018</td>
<td>-0.26**</td>
<td>0.044</td>
<td>0.355</td>
<td>-0.18**</td>
<td>0.031</td>
<td>-0.07*</td>
<td>0.000</td>
<td>-0.04**</td>
<td>-0.04**</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.05 level (2-tailed).
**Correlation is significant at the 0.01 level (2-tailed).
## Correlations: Rice Sample

<table>
<thead>
<tr>
<th>Gender</th>
<th>ETHNIC2</th>
<th>SAT</th>
<th>TIME</th>
<th>COMPLEX</th>
<th>CONTRA</th>
<th>VAGUE</th>
<th>SIZE</th>
<th>GREEN</th>
<th>ALTGREEN</th>
<th>ERSMINUS</th>
<th>RATEMEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sig. (2-tailed)</td>
<td>1</td>
<td>.009</td>
<td>-.063</td>
<td>.056</td>
<td>-.013</td>
<td>-.058</td>
<td>.053</td>
<td>-.231**</td>
<td>.066</td>
<td>.175*</td>
<td>144</td>
</tr>
<tr>
<td>N</td>
<td>127</td>
<td>127</td>
<td>115</td>
<td>122</td>
<td>127</td>
<td>126</td>
<td>126</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>125</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.009</td>
<td>1</td>
<td>-.143</td>
<td>.019</td>
<td>.063</td>
<td>.134</td>
<td>.136</td>
<td>.015</td>
<td>-.112</td>
<td>.115</td>
<td>.015</td>
</tr>
<tr>
<td>N</td>
<td>127</td>
<td>127</td>
<td>115</td>
<td>122</td>
<td>127</td>
<td>126</td>
<td>126</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>125</td>
</tr>
<tr>
<td>SAT</td>
<td>Pearson Correlation</td>
<td>-.053</td>
<td>-.143</td>
<td>.1</td>
<td>.132</td>
<td>-.301**</td>
<td>-.277*</td>
<td>-.271**</td>
<td>-.213*</td>
<td>-.040</td>
<td>-.132</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.503</td>
<td>.127</td>
<td>.170</td>
<td>.001</td>
<td>.003</td>
<td>.003</td>
<td>.047</td>
<td>.047</td>
<td>.024</td>
<td>.694</td>
<td>.151</td>
</tr>
<tr>
<td>N</td>
<td>115</td>
<td>115</td>
<td>115</td>
<td>110</td>
<td>151</td>
<td>115</td>
<td>115</td>
<td>115</td>
<td>115</td>
<td>115</td>
<td>114</td>
</tr>
<tr>
<td>TIME</td>
<td>Pearson Correlation</td>
<td>.056</td>
<td>.019</td>
<td>-.132</td>
<td>.1</td>
<td>-.234**</td>
<td>-.145</td>
<td>-.201*</td>
<td>-.168*</td>
<td>.267*</td>
<td>.152</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.523</td>
<td>.835</td>
<td>.170</td>
<td>.009</td>
<td>.112</td>
<td>.027</td>
<td>.047</td>
<td>.047</td>
<td>.024</td>
<td>.694</td>
<td>.151</td>
</tr>
<tr>
<td>N</td>
<td>122</td>
<td>122</td>
<td>110</td>
<td>122</td>
<td>122</td>
<td>121</td>
<td>121</td>
<td>122</td>
<td>122</td>
<td>122</td>
<td>120</td>
</tr>
<tr>
<td>COMPLEX</td>
<td>Pearson Correlation</td>
<td>-.013</td>
<td>.063</td>
<td>-.301**</td>
<td>-.234**</td>
<td>1</td>
<td>.804**</td>
<td>.508**</td>
<td>.162</td>
<td>-.139</td>
<td>-.285**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.883</td>
<td>.356</td>
<td>.001</td>
<td>.009</td>
<td>.000</td>
<td>.000</td>
<td>.007</td>
<td>.123</td>
<td>.066</td>
<td>.001</td>
<td>.452</td>
</tr>
<tr>
<td>N</td>
<td>127</td>
<td>127</td>
<td>115</td>
<td>122</td>
<td>127</td>
<td>126</td>
<td>126</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>125</td>
</tr>
<tr>
<td>CONTRA</td>
<td>Pearson Correlation</td>
<td>-.058</td>
<td>.134</td>
<td>-.277*</td>
<td>-.145</td>
<td>.004**</td>
<td>1</td>
<td>.500**</td>
<td>.216*</td>
<td>-.039</td>
<td>-.638</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.518</td>
<td>.136</td>
<td>.003</td>
<td>.112</td>
<td>.000</td>
<td>.000</td>
<td>.015</td>
<td>.065</td>
<td>.069</td>
<td>.029</td>
<td>.890</td>
</tr>
<tr>
<td>N</td>
<td>126</td>
<td>126</td>
<td>115</td>
<td>121</td>
<td>126</td>
<td>126</td>
<td>126</td>
<td>126</td>
<td>126</td>
<td>126</td>
<td>125</td>
</tr>
<tr>
<td>VAGUE</td>
<td>Pearson Correlation</td>
<td>.063</td>
<td>.136</td>
<td>-.271**</td>
<td>-.201*</td>
<td>.566**</td>
<td>.550**</td>
<td>.1</td>
<td>.124</td>
<td>.042</td>
<td>.066</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.481</td>
<td>.130</td>
<td>.003</td>
<td>.027</td>
<td>.000</td>
<td>.000</td>
<td>.168</td>
<td>.463</td>
<td>.051</td>
<td>.035</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>126</td>
<td>126</td>
<td>115</td>
<td>121</td>
<td>126</td>
<td>126</td>
<td>126</td>
<td>126</td>
<td>126</td>
<td>126</td>
<td>125</td>
</tr>
<tr>
<td>SIZE</td>
<td>Pearson Correlation</td>
<td>-.231**</td>
<td>.015</td>
<td>-.213*</td>
<td>-.180*</td>
<td>.162</td>
<td>.216*</td>
<td>.124</td>
<td>1</td>
<td>-.062</td>
<td>-.084</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.009</td>
<td>.865</td>
<td>.022</td>
<td>.047</td>
<td>.070</td>
<td>.015</td>
<td>1</td>
<td>.496</td>
<td>.350</td>
<td>.442</td>
<td>.018</td>
</tr>
<tr>
<td>N</td>
<td>127</td>
<td>127</td>
<td>115</td>
<td>122</td>
<td>127</td>
<td>126</td>
<td>126</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>125</td>
</tr>
<tr>
<td>GREEN</td>
<td>Pearson Correlation</td>
<td>.968</td>
<td>-.122</td>
<td>-.040</td>
<td>.207*</td>
<td>-.139</td>
<td>-.038</td>
<td>.042</td>
<td>-.002</td>
<td>1</td>
<td>.513**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.453</td>
<td>.215</td>
<td>.671</td>
<td>.024</td>
<td>.122</td>
<td>.605</td>
<td>.645</td>
<td>.406</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>124</td>
<td>124</td>
<td>113</td>
<td>119</td>
<td>124</td>
<td>124</td>
<td>124</td>
<td>124</td>
<td>124</td>
<td>124</td>
<td>124</td>
</tr>
<tr>
<td>ALTGREEN</td>
<td>Pearson Correlation</td>
<td>.175*</td>
<td>.115</td>
<td>-.132</td>
<td>.152</td>
<td>-.153</td>
<td>-.038</td>
<td>.096</td>
<td>.084</td>
<td>.513**</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.049</td>
<td>.199</td>
<td>.159</td>
<td>.094</td>
<td>.086</td>
<td>.086</td>
<td>.463</td>
<td>.350</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>127</td>
<td>127</td>
<td>115</td>
<td>122</td>
<td>127</td>
<td>126</td>
<td>126</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>125</td>
</tr>
<tr>
<td>ERSMINUS</td>
<td>Pearson Correlation</td>
<td>.144</td>
<td>-.015</td>
<td>.042</td>
<td>.131</td>
<td>-.285*</td>
<td>-.195*</td>
<td>.041</td>
<td>.041</td>
<td>.069</td>
<td>.588**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.107</td>
<td>.867</td>
<td>.857</td>
<td>.151</td>
<td>.001</td>
<td>.029</td>
<td>.651</td>
<td>.442</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>127</td>
<td>127</td>
<td>115</td>
<td>122</td>
<td>127</td>
<td>126</td>
<td>126</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>125</td>
</tr>
<tr>
<td>RATEMEAN</td>
<td>Pearson Correlation</td>
<td>-.062</td>
<td>-.044</td>
<td>-.146</td>
<td>-.135</td>
<td>-.038</td>
<td>.012</td>
<td>.189</td>
<td>.496</td>
<td>.211*</td>
<td>.375**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.365</td>
<td>.625</td>
<td>.121</td>
<td>.141</td>
<td>.452</td>
<td>.890</td>
<td>.035</td>
<td>.018</td>
<td>.000</td>
<td>.131</td>
<td>.038</td>
</tr>
<tr>
<td>N</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
<td>125</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).