2017

Patriarchy, Not Hierarchy: Rethinking the Effect of Cultural Attitudes in Acquaintance Rape Cases

Eric R. Carpenter

Florida International University College of Law, ercarpen@fiu.edu

PATRIARCHY, NOT HIERARCHY: RETHINKING THE EFFECT OF CULTURAL ATTITUDES IN ACQUAINTANCE RAPE CASES

Eric R. Carpenter[*]

ABSTRACT

Do certain people view acquaintance rape cases in ways that favor the man? The answer to that question is important. If certain people do, and those people form a disproportionately large percentage of the people in the institutions that process these cases, then those institutions may process these cases in ways that favor the man.

In 2010, Dan Kahan published Culture, Cognition, and Consent, *a study on how people evaluate a dorm room rape scenario. He found that those who endorsed a stratified, hierarchical social order were more likely to find that the man should not be found guilty of rape.*

If Kahan is right, radical change may be necessary. The institutions responsible for handling sexual assault complaints – law enforcement communities, the military, and university and college administrations – are stratified and hierarchical, and are likely over-populated by people who are attracted to hierarchical institutions and who hold hierarchical world views. These institutions may need to be overhauled – or even replaced.

However, the study has a serious methodological flaw: it uses the Hierarchy-Egalitarianism Scale to measure those hierarchical world views, and as this article demonstrates, this scale has reliability and validity issues.

This article then applies a different methodology to the underlying data and shows that patriarchy, not hierarchy, explains the differences in guilt perceptions. This more accurate understanding of Kahan's data carries important policy implications. Rather than radical change, targeted training that addresses inaccurate rape beliefs may be enough to ensure accurate processing of these cases.

---

[*] Assistant Professor of Law, FIU College of Law. I thank Dan Kahan for sharing his data and Ben Fay for his assistance with the statistics. I also thank Mae Quinn, Paul Gowder, Barbara O'Brien, Alex Acosta, Ediberto Román, and Howard Wasserman for reviewing earlier drafts. This article benefited from a presentation at the Southeastern Association of Law Schools Annual Conference.

Table of Contents

I. INTRODUCTION

Do certain people view acquaintance rape cases in ways that favor the man? The answer to that question is important. If certain people do, and those people form a disproportionately large percentage of the people in the institutions that are responsible for handling sexual assault cases, then those institutions might process these cases in ways that favor the man.

And if what characterizes this group is a deeply held belief system, then radical change may be necessary. Trying to get someone to change deeply-held beliefs would be akin to trying to change someone from conservative to liberal, or Christian to Jew. The institutions themselves may need to be overhauled – or even replaced.

In 2010, Dan Kahan published *Culture, Cognition, and Consent*,[1] which included an important finding: those who endorsed a stratified, hierarchical social order were more likely to side with the man and find him not guilty of a dorm room rape scenario.[2] This finding informs the problem described above. The institutions responsible for handling sexual assault complaints – law enforcement communities, the military, and university and college administrations – are stratified and hierarchical, and are likely populated by people who are attracted to hierarchical institutions and who hold hierarchical world views. If Kahan is right, radical change may be necessary.

Unfortunately, his study may have a serious methodological problem. The study used the Hierarchy-Egalitarianism Scale (H-ES), which was designed by the Cultural Cognition Project (CCP)[3] to measure those hierarchical world views. This article shows that the scale has reliability issues (because of ambiguous and complex items) and validity issues (because it may not be measuring hierarchy). As a result of these problems, the link between hierarchical worldviews and perceptions of guilt in sexual-assault cases may not be accurate.

Kahan's underlying data remains valuable, though. Most research on sexual assault is conducted on a small population – college students. Here, Kahan gathered data from a large sample that is representative of the general

---

[1] Dan M. Kahan, *Culture, Cognition, and Consent: Who Perceives What, and Why, in Acquaintance-Rape Cases*, 158 U. PA. L. REV. 729 (2010) [hereinafter *Culture, Cognition, and Consent*].

[2] *Id.* at 733, 776, 793-94.

[3] The CCP is "a group of scholars interested in studying how cultural values shape public risk perceptions and related policy beliefs". Cultural Cognition Project, http://www.culturalcognition.net/ (last visited Jan. 31, 2016). Dan Kahan is a member of this group. The scale that I will be analyzing in this article was developed by this group rather than by one person. Various members of this group, separately and jointly, have published articles using this study. For clarity, in the main text I will attribute the development of the H-ES to the CCP and in the footnotes I will cite to the particular authors.

population. Further, the study uses a realist hypothetical of a dorm-room sexual assault, and this is the type of case that is at the center of the sexual assault controversy.

This article uses a different analysis of the items used in the CCP's scale to show that patriarchy, rather than hierarchy, predicts how people view sexual assault cases. This more accurate understanding of Kahan's data carries important policy implications. If beliefs about patriarchy account for the variation in guilt judgments in these cases, then it may not be necessary to radically restructure institutions. There is nothing inherently wrong with hierarchical organizations handling sexual-assault cases. Instead, targeted training related to these more narrow belief systems about gender roles may be sufficient to ensure more accurate processing of these cases.

In Part II, this article reviews Kahan's study. In Part III, this article analyzes the H-ES. In Part IV, this article offers Kahan's argument that the H-ES is unidimensional and valid within *Culture, Cognition, and Consent*. In Part V, this article conducts exploratory factor analysis of the H-ES that suggests that the scale is multi-dimensional. In Part VI, this article conducts structural equation modeling with a theoretical model that has gender and race facets and where the gender facet predicts the guilt variable while the race facet does not. This Part also shows that the CCP's single construct model has a poor fit to the data. In Part VII, this article discusses the impact of these findings on other CCP studies and suggest improvements to the scale. In Part VIII, this article concludes with the policy implications.

II. KAHAN'S STUDY AND FINDINGS

In 2009, Kahan conducted an experiment to see whether people who subscribe to certain worldviews perceive rape cases differently than people with contrasting worldviews. The respondents (n = 1,500) were given a vignette of a sexual assault involving Lucy and Dave, two college students and casual acquaintances. One day, when Lucy was looking for her boyfriend in the dorms, she stopped by Dave's room to see his roommate. She had had a drink prior to the visit. She went into the dorm room but the roommate was not there; however, Dave was. At this point, Lucy claims that she tried to leave but Dave blocked the door, pinned her down, and sexually assaulted her by inserting his penis into her vagina. Dave claims that she consented. Both agree that during the event, Lucy said "no" repeatedly (although Dave claims that she said it in a sexual way) and both agree that Lucy did not otherwise physically resist.

The CCP randomly divided the subjects into five groups of 300 and gave each of them one of five legal conditions (either no law to apply, or four different versions of law to apply) which the respondents would use to

evaluate the facts in the case.[4] Kahan measured thirteen dependent variables, to include an important outcome judgment, guilt ("Dave should be found guilty of rape.").

For use as independent variables, Kahan gathered demographic information from the respondents and also administered the H-ES.[5] He coded respondents as "hierarchs" if the respondent scored in the top third of the scale and as "egalitarian" if the respondent scored in the bottom third. He then used "hierarch" and "egalitarian" as predictor variables.[6]

Kahan found that hierarchs were more likely than egalitarians to side with the man and find him not guilty of rape.[7] He also found that gender only had a meaningful effect when joined with the cultural world view,[8] such that hierarchical women were actually the most predisposed to side with the man.[9]

This study is not inconsistent with other research, but it does point to a more global variable as the explanatory variable. Other research has pointed to a more discrete variable: traditional, hierarchical gender role beliefs.[10] Those with traditional gender role beliefs tend to endorse certain inaccurate rape schemas more than those with non-traditional gender role beliefs. These beliefs are then associated with more discrete beliefs about rape, and ultimately with the outcome judgments in rape hypotheticals.[11] Studies have

---

[4] *Culture, Cognition, and Consent*, *supra* note 1, at 767-68.

[5] The CCP also administered another scale, the Individualism-Communitarian Scale (I-CS). *See infra* Part IV.B.

[6] *Culture, Cognition, and Consent*, *supra* note 1, at 776 fig.3, 777 fig.4, 780 fig.5, 785 fig.7, 786 fig.8, 792 fig.10.

[7] *Id.* at 733, 776, 793-94.

[8] *Id.* at 782.

[9] *Id.* at 787.

[10] The traditional gender role construct has many potential facets: beliefs that the man should be in charge of the family unit; that women should remain at home rather than work outside the home; that men should pursue women while women should be passive; and that women should behave in sexually conservative ways. *See* Carpenter, *supra* note X, at 391-92. Of these facets, it is likely that expectations about sexual conservatism (particularly, that women should be lady-like) is the facet that plays the most central role in rape case processing. *Id*. at 394.

[11] *See* Dominic Abrams et al., *Perceptions of Stranger and Acquaintance Rape: The Role of Benevolent and Hostile Sexism in Victim Blame and Rape Proclivity*, 84 J. PERSONALITY & SOC. PSYCHOL. 111 (2003); Kathryn B. Anderson et al., *Individual Differences and Attitudes Toward Rape: A Meta-Analytic Review*, 23 PERSONALITY & SOC. PSYCHOL. BULL. 295, 312 (1997); Gordon B. Forbes et al., *First- and Second-Generation Measures of Sexism, Rape Myths and Related Beliefs, and Hostility Toward Women*, 10 VIOLENCE AGAINST WOMEN 236, 250 (2004); Barbara E. Johnson et al., *Rape Myth Acceptance and Sociodemographic Characteristics: A Multidimensional Analysis*, 36 SEX ROLES 693, 704 (1997); Laura L. King & Jennifer J. Roberts, *Traditional Gender Role and Rape Myth Acceptance: From the Countryside to the Big City*, 21 WOMEN & CRIM. JUST. 1, 9, 12 (2011); Eliana Suarez & Tahany M. Gadalla, *Stop Blaming the Victim: A Meta-Analysis on Rape Myths*, 25 J. INTERPERSONAL VIOLENCE 2010, 2022 (2010); Lynda A. Szymanski

also found that acceptance of these inaccurate rape schemas is associated with siding with the man in the ultimate normative judgment about blame.[12]

However, Kahan's findings are only valid if the H-ES is itself valid and reliable. If the H-ES actually measures something else or does not measure anything at all, then those findings may be inaccurate and policy makers should not rely on them when making decisions. The next section turns to that issue – whether the H-ES is a valid and reliable scale.

III. ANALYSIS OF THE HIERARCHY-EGALITARIANISM SCALE

A. Overview of Scale Development

Researchers generally follow certain steps when developing scales, and these steps provide a useful framework for evaluating the H-ES. The scale development steps are: 1) clearly define the construct; 2) generate the item pool; 3) determine format for measurement; 4) have the item pool reviewed by a panel of experts; 5) consider inclusion of validation items; 6) administer items to a development sample; 7) evaluate the items; and 8) decide on scale length.[13] When followed, these steps help to ensure that the resulting scale is reliable and valid.

Reliability "is the extent to which it is possible to replicate a measurement, reproducing the same value (regardless of whether it is the right one) on the same standard for the same subject at the same time."[14] As Lee Epstein and Gary King explain this concept, "If any one of us stepped on the same bathroom scale one hundred times in a row, and if the scale were working reliably, it would give us the same weight one hundred times in a row – even if that weight were not accurate."[15]

---

et al., *Gender Role and Attitudes Toward Rape in Male and Female College Students*, 29 SEX ROLES 37 (1993); G. Tendayi Viki & Dominic Abrams, *But She Was Unfaithful: Benevolent Sexism and Reactions to Rape Victims Who Violate Traditional Gender Role Expectations*, 47 SEX ROLES 289 (2002); Rosanne Proite et al., *Gender, Sex-role Stereotypes, and the Attribution of Responsibility for Date and Acquaintance Rape*, 34 J. C. STUDENT DEV. 411 (1993); Niwako Yamawaki, *Rape Perception and the Function of Ambivalent Sexism and Gender-Role Traditionality*, 22 J. INTERPERSONAL VIOLENCE 406 (2007).

[12] *See, e.g.*, Barbara Krahe, *Social Psychological Issues in the Study of Rape*, 2 EUR. SOC. PSYCHOL. 279 (1991); G. Tendayi Viki et al., *Evaluating Stranger and Acquaintance Rape: The Role of Benevolent Sexism in Perpetrator Blame and Recommended Sentence Length*, 28 L. & HUM. BEHAV. 295 (2004); Charlene Muehlenhard, *Misinterpreting Dating Behaviors and the Risk of Date Rape*, 6 J. SOC. & CLINICAL PSYCHOL. 20 (1988).

[13] ROBERT F. DEVELLIS, SCALE DEVELOPMENT: THEORY AND APPLICATIONS ch. 5 (3d ed. 2012). *See also* Lee Anna Clark & David Watson, *Constructing Validity: Basic Issues in Objective Scale Development*, 7 PSYCHOL. ASSESSMENT 309 (1995).

[14] Lee Epstein & Gary King, *The Rules of Inference*, 69 U. CHI. L. REV. 1, 83 (2002).

[15] *Id.*

Researchers who are studying a latent, unobservable variable often use a scale that consists of several questions that can measure discrete items that are related to that unobservable variable, and then use a score based on those items (called a factor score) to represent the value of that unobservable variable. There, "reliability concerns how much a variable influences a set of items"[16] and "[t]he more the score we obtain from a scale represents the true score of the [unobserved] variable and the less it reflects other extraneous factors, the more reliable our scale is."[17]

Validity, on the other hand, "is the extent to which a reliable measure reflects the underlying concept being measured."[18] Returning to Epstein and King's scale example, "If one's true weight is 150 and the scale, even one hundred times in a row, reports 125, we would not think much of that scale."[19] That scale would be reliable but not valid.

In this study, the specific issue is whether the items in the H-ES measure the CCP's targeted global construct or whether instead the items measure another construct, like patriarchy, or even no construct at all. This is an issue of construct validity, which "is the degree to which an assessment instrument measures the targeted construct."[20] A component of construct validity is content validity. Content validity is "the degree to which elements of an assessment instrument are relevant to and representative of the targeted construct for a particular assessment purpose."[21]

Scale developers generally address content validity with steps 1, 2, and 4. Sale developers clearly define the construct; generate an item pool; and then have that item pool reviewed by experts. The CCP has not published a detailed article on the H-ES' psychometric properties so what follows comes from various sources.

B. The CCP's Construct Definition: "Grid"

Because "[c]ontent validity is intimately linked to the definition of the construct being examined",[22] we need to identify precisely what the CCP was

---

[16] DEVELLIS, *supra* note 13, at 59.

[17] *Id.* at 31.

[18] Epstein & King, *supra* note 14, at 87.

[19] *Id.*

[20] Stephen N. Haynes et al., *Content Validity in Psychological Assessment: A Functional Approach to Concepts and Methods*, 7 PSYCHOL. ASSESSMENT 238, 239 (1995).

[21] Haynes et al., *supra* note 20, at 238, 239. *See also* DEVELLIS, *supra* note 13, at 64 (construct validity is "the extent to which a measures 'behaves' the way the construct it purports to measure should behave with regard to established measures of other constructs.") DeVellis contrasts this to criterion validity, which is the ability of a scale to predict relationships among variables. *Id.* at 61-62.

[22] DEVELLIS, *supra* note 13, at 60, 73; Clark & Watson, *supra* note 13, at 310.

trying to measure. The CCP's construct is derived from the cultural theory of risk developed by Mary Douglas, Michael Thompson, Aaron Wildavsky, and Karl Dake.[23] The cultural theory of risk proposes that "individuals select certain risks for attention and disregard others in a way that reflects and reinforces the particular worldviews to which they adhere."[24]

The theory advances a "grid/group" taxonomy. The "grid" dimension captures group regulation, social prescription, and structured social order. The dimension "runs from minimum to maximum regulation",[25] where "[i]n a high-grid environment, everything is classified and individual choice is heavily restricted."[26] There, structured social roles regulate the actions between individuals. Low-grid environments lack structure and "individuals are increasingly expected to negotiate their own relationships with others."[27]

The "group" dimension captures group integration, identity, and the general boundary around that community:[28] it is "the amount of moral pressure to conform that a community puts on its members."[29] The higher the "group" value, "the tighter the control over admission into the group and the higher the boundaries separating members from nonmembers."[30]

Four cultures coalesce at the extremes of those dimensions: hierarchs, egalitarians, individualists, and fatalists:[31]

---

[23] Dan M. Kahan & Donald Braman, *Cultural Cognition and Public Policy*, 24 YALE L. & POL'Y REV. 149, 152, 157 (2006) [hereinafter *Cultural Cognition and Public Policy*].

[24] *Cultural Cognition and Public Policy*, *supra* note 23, at 154.

[25] Mary Douglas, *Being Fair to Hierarchists*, 151 U. PA. L. REV. 1349, 1353 (2003) [hereinafter *Being Fair to Hierarchists*] at 1353.

[26] *Id.* at 1352.

[27] MICHAEL THOMPSON ET AL., CULTURAL THEORY 6 (1990).

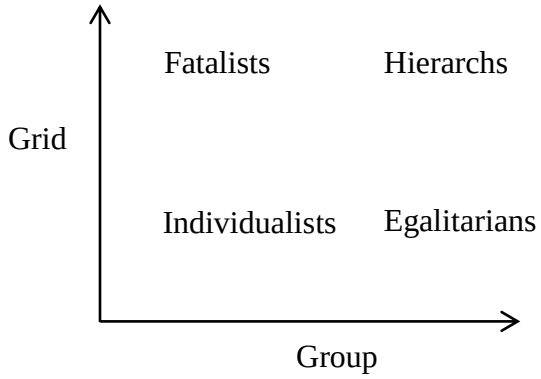
[28] Karl Dake, *Myths of Nature: Culture and the Social Construction of Risk*, 48 J. SOC. ISSUES 21, 28 (1992) [hereinafter, *Myths of Nature*]; *Being Fair to Hierarchists*, *supra* note 25, at 1355.

[29] *Being Fair to Hierarchists*, *supra* note 25, at 1352.

[30] THOMPSON ET AL., *supra* note 27, at 6.

[31] For a detailed discussion of these cultural categories, *see* THOMPSON ET AL., *supra* note 27, at 5-11; *see generally* MARY DOUGLAS & AARON WILDAVSKY, RISK AND CULTURE (1982); Aaron Wildavsky & Karl Dake, *Theories of Risk Perception: Who Fears What and Why?*, 119 DAEDALUS 41 (1990); Aaron Wildavsky, *Choosing Preferences by Constructing Institutions: A Cultural Theory of Preference Formation*, 81 AM. POL. SCI. REV. 3, 6 (1987); *Myths of Nature*, *supra* note 28, at 28-30.

Figure 1: Douglas typology



The "prime virtue" for hierarchists is order; for egalitarians, justice; and for individualists, liberty.[32] Fatalists are drop-outs or cast-outs, politically inactive and apathetic.[33] These four cultures are competitive and need their adversarial relationships to help define their own legitimacy.[34]

The basis for the taxonomy is "the competing cultures of hierarchism and individualism."[35] They sit on the opposite, extreme parts of the taxonomy, where hierarchs are high-group and high-grid, and individualists are low-group and low-grid.[36] Hierarchs and egalitarians both have strong group boundaries, but hierarchs differ from egalitarians in that hierarchs also have strong internal boundaries and classifications while egalitarians are free to do what they like within the group with minimal regulation.[37] While hierarchs are set opposite to individualists, egalitarians are most strongly contrasted against hierarchs because they are the organized dissidents of that authoritative and regulated culture.[38]

While "grid" and "group" provide the taxonomy for the four worldviews, "grid" and "group" are not the actual constructs that these earlier scholars tried to measure. For them, the worldviews are the constructs. Dake and Wildavsky developed scales to measure them,[39] eventually focusing on just three (egalitarian, hierarch, and individualist, and ignoring fatalist).[40] Other

---

[32] *Being Fair to Hierarchists*, *supra* note 25, at 1367.

[33] *Id.* at 1369.

[34] Mary Douglas, *Four Cultures: The Evolution of a Parsimonious Model*, 47 GEOJOURNAL 411, 413 (1999).

[35] *Being Fair to Hierarchists*, *supra* note 25, at 1353.

[36] *Id.* at 1352.

[37] *Id.* at 1353.

[38] *Id.* at 1368-69.

[39] *Myths of Nature*, *supra* note 28, at 31.

[40] Karl Dake & Aaron Wildavsky, *Individual Differences in Risk Perception and Risk-Taking Preferences*, *in* THE ANALYSIS, COMMUNICATION, AND PERCEPTION OF RISK 15, 21

researchers have followed this model and have used similar scales.[41]

The CCP scholars connect aspects of cultural theory with similar features found in social psychology, arguing that "cultural commitments operate as a kind of heuristic in the rational processing of information on public policy."[42] They argue that individual citizens cannot personally investigate the facts that are used to support public policy arguments; therefore, citizens have to choose whom to trust among those who are advancing "facts" – and citizens trust those who share their worldviews.[43] This is congruent with social identity theory[44] and recognized psychological principles that explain motivated reasoning, like confirmation bias and the avoidance of cognitive dissonance.[45] Under social identity theory, those in the "ingroup" trust others in the ingroup and disbelieve those in the "outgroup." The CCP uses the cultural worldviews to define these ingroups and outgroups.

From this flows the CCP's definition of cultural cognition: "Cultural cognition refers to the tendency of individuals to conform their beliefs about disputed matters of fact (e.g., whether global warming is a serious threat; whether the death penalty deters murder; whether gun control makes society more safe or less) to values that define their cultural identities."[46] The CCP

(B.J. Garrick & W.C. Gekler, eds., 1991); Karl Dake, *Orienting Dispositions in the Perception of Risk: An Analysis of Contemporary Worldviews and Cultural Biases*, 22 J. CROSS-CULTURAL PSYCHOL. 61, 65 (1991); Wildavsky & Dake, *supra* note 31, at 47.

[41] Hank C. Jenkins-Smith & Walter K. Smith, *Ideology, Culture, and Risk Perception*, *in* POLITICS, POLICY, AND CULTURE 17, 22 (Dennis J. Coyle & Richard J. Ellis eds., 1994) (not using fatalist); Ellen Peters & Paul Slovic, *The Role of Affect and Worldviews as Orienting Dispositions in the Perception and Acceptance of Nuclear Power*, 26 J. APPLIED SOC. PSYCHOL. 1427, 1433-35 (1998) (combining fatalist and hierarchist); Jean Brenot et al., *Testing the Cultural Theory of Risk in France*, 18 RISK ANALYSIS 729, 732 (1998) (using all four constructs); Richard J. Ellis & Fred Thompson, *Seeing Green: Cultural Biases and Environmental Preferences*, *in* CULTURAL MATTERS: ESSAYS IN HONOR OF AARON WILDAVSKY 169, 182 (Richard J. Ellis & Michael Thompson eds., 1997) (not using fatalist); Claire Marris et al., *A Quantitative Test of the Cultural Theory of Risk Perceptions: Comparison with the Psychometric Paradigm*, 18 RISK ANALYSIS 635, 638 (1998) (using all four constructs). Mary Douglas thinks that ignoring fatalists is a mistake. *Being Fair to Hierarchists*, *supra* note 25, at 1369-70.

[42] *Cultural Cognition and Public Policy*, *supra* note 23, at 151; *see also* Dan M. Kahan, *Cultural Cognition as a Conception of the Cultural Theory of Risk*, in HANDBOOK OF RISK THEORY 725 (Sabine Roeser et al. eds. 2012) [hereinafter, HANDBOOK OF RISK THEORY).

[43] *Cultural Cognition and Public Policy*, *supra* note 23, at 151.

[44] *See generally* Matthew J. Hornsey, *Social Identity Theory and Self-categorization Theory: A Historical Review*, 2 SOC. & PERSONALITY PSYCHOL. COMPASS 204 (2008).

[45] *See generally Cultural Cognition and Public Policy*, *supra* note 23, at 165-66; Donald Braman et al., *Modeling Facts, Culture, and Cognition in the Gun Debate*, 18 SOC. JUST. RES. 283, 289 (2005); HANDBOOK OF RISK THEORY, *supra* note 42, at 739. *See generally* Cass R. Sunstein, MISFEARING: A REPLY, 119 HARV. L. REV. 1110 (2006).

[46] CULTURAL COGNITION PROJECT, http://www.culturalcognition.net/ (last visited Jan. 31, 2016).

argues that these worldviews, rather than other predictive variables like race and sex, best explain how people perceive risk.[47]

The CCP then modified the cultural theory taxonomy. For the purposes of this project, this move is important because it identifies the constructs that the CCP was trying to measure. The CCP continued to have "group" on the x-axis and "grid" on the y-axis and used three of the terms that other researchers continued to focus on – hierarchy, egalitarianism, and individualism.[48] However, the CCP added a new term, communitarianism,[49] and set individualism and communitarianism opposite each other on the x-axis to represent poles of the "group" dimension.[50] This move is generally consistent with the Douglas "group" dimension, where the "individualism" label represents low group identity, and the "communitarianism" label represents high group identity.

The CCP kept "grid" on the y-axis but then set hierarchy and egalitarianism opposite each other on this axis, and importantly, this move is inconsistent with the Douglas model. Under the Douglas model, these two worldviews were the products of the two dimensions of "grid" and "group." They were not the poles of the "grid" dimension. For Douglas, the poles of the "grid" dimension would have been "structure" and "structureless."

The CCP then renamed the type of people that would fall within the resulting quadrants.[51] In the quadrant where Wildavsky and Dake set a

---

[47] *Cultural Cognition and Public Policy*, *supra* note 23, at 158-59. *See generally Culture, Cognition, and Consent*, *supra* note 1, at 733-34; Dan M. Kahan et al., *Gender, Race, and Risk Perception: The Influence of Cultural Status Anxiety* (Apr. 7, 2005) (unpublished manuscript), available at http:ssrn.com/abstract=723762 [hereinafter, *Gender, Race, and Risk Perception*] (this article was later revised and published as Dan M. Kahan et al., *Culture and Identity-Protective Cognition: Explaining the White-Male Effect in Risk Perception*, 4 J. EMPIRICAL LEGAL STUD. 465 (2007) [hereinafter *Culture and Identity-Protective Cognition*]); HANDBOOK OF RISK THEORY, *supra* note 42, at 741.

[48] Dan M. Kahan & Donald Braman, *More Statistics, Less Persuasion: A Cultural Theory of Gun-Risk Perception*, 151 U. PA. L. REV. 1291, 1297 (2003) [hereinafter *More Statistics, Less Persuasion*], *citing* Wildavsky, *supra* note 31, at 11-13; *see also* Wildavsky & Dake, *supra* note 3, at 44.

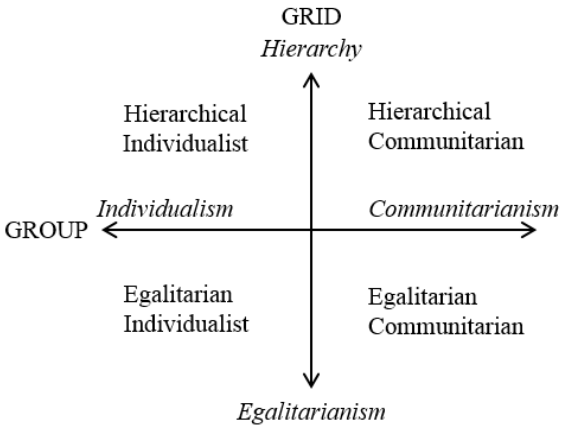
[49] *More Statistics, Less Persuasion*, *supra* note 48, at 1303. The CCP initially used the term "solidarism" but later began using "communitarianism" instead. HANDBOOK OF RISK THEORY, *supra* note 42, at 730.

[50] *Cultural Cognition and Public Policy*, *supra* note 23, at 153.

[51] *See, e.g*., Dan M. Kahan et al., *Cultural Cognition of Scientific Consensus*, 14 J. RISK RES. 147, 150 (2011) [hereinafter, *Cultural Cognition of Scientific Consensus*]; Dan M. Kahan et al., *Affect, Values, and Nanotechnology Risk Perceptions: An Experimental Investigation* 11 (Yale Law Sch. Pub. Law Working Paper, Paper No. 155 (2007)) [hereinafter, *Affect, Values, and Nanotechnology Risk Perceptions*] (this study also formed the basis for Dan M. Kahan et al., *The Future of Nanotechnology Risk Perceptions: An Experimental Investigation of Two Hypothesis* (Cultural Cognition Working Paper, Paper No. 46 (2008)); HANDBOOK OF RISK THEORY, *supra* note 42, at 739.

worldview of "hierarchy," the CCP placed the label "hierarchical communitarian;" "egalitarianism" became "egalitarian communitarian;" "individualism" became "egalitarian individualist;" and "fatalist" became "hierarchical individualist." Note that while Wildavsky and Dake posited four separate constructs based on worldviews and created scales to measure those constructs, the CCP now posited just two constructs – grid and group.[52]

Figure 2: CCP typology



The CCP provides definitions of these two constructs. "Grid" represents a single, continuous measure of "how favorably or unfavorably disposed individuals are toward a social order that features differentiation and stratification of social roles based on observable and largely fixed characteristics (including race, gender, sexual orientation, and class)."[53] At

[52] HANDBOOK OF RISK THEORY, *supra* note 42, at 730. The CCP often refers to the "grid" construct as "hierarchy."

[53] *Culture, Cognition, and Consent*, *supra* note 1, at 770. The "observable and largely fixed characteristics" part of the definition represents a significant departure from the Douglas typology. For a criticism of this modification, *see Being Fair to Hierarchists*, *supra* note 25, at 1364. Douglas believed that the CCP construct equated hierarchs to racists and sexists. Id. at 1362-63. For the CCP response, *see* Dan M. Kahan & Donald Braman, *Caught in the Crossfire: A Defense of the Cultural Theory of Gun-Risk Perceptions*, 151 U. PA. L. REV. 1395, 1408 (2003) [hereinafter, *Caught in the Crossfire*]; *see also* HANDBOOK OF RISK THEORY, *supra* note 42. It may be that the CCP constructs have departed significantly from original cultural theory constructs and their constructs now more closely resemble those associated with the Social Dominance Orientation Scale, Felicia Pratto et al., *Social Dominance Orientation: A Personality Variable Predicting Social and Political Attitudes*, 67 J. PERSONALITY AND SOC. PSYCHOL. 741, 760 (1994), or the Right-Wing Authoritarianism Scale, Bob Altemeyer, *The Other "Authoritarian Personality"*, *in* 30 ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY 85, 86-87 (1998). *See also* Bernard E. Whitley, Jr., *Right-Wing Authoritarianism, Social Dominance Orientation, and Prejudice*, 77 J. PERSONALITY AND SOC. PSYCHOL. 126 (1999). The items from these scales are

the poles, the CCP defines "hierarchy" as "favors deference to traditional forms of social and political authority and is protective of the roles and status claims they entail"[54] and where "entitlements, obligations, opportunities and offices are all assigned on the basis of conspicuous and largely fixed attributes, such as gender, race, lineage, class, and the like".[55] The CCP defines "egalitarian" as "abhors social stratification, distrusts the social and political authority structures that rest on such differentiation, and favors collective action to equalize wealth, status, and power".[56]

"Group" represents a single, continuous measure of "how favorably or unfavorably disposed individuals are toward a social order that treats individuals as responsible for securing the conditions of their own flourishing without collective assistance and that resists collective interference with individual strivings."[57] The poles of this attitude are individualism and collectivism. The CCP defines "individualist" as "prizes individual autonomy, celebrates free markets and other institutionalized forms of private ordering, and resents collective interference with the same"[58] and where individuals "are expected to secure the conditions of their own flourishing without interference or assistance from the collective"[59] The CCP defined "communitarianism" (or "solidarism") as "logically opposed to individualism"[60] and where "collective needs trump individual initiative, and in which society is expected to secure the conditions of individual flourishing."[61]

The CCP created two scales to measure those constructs, measuring the grid construct with the H-ES and the group construct with the Individualism-Communitarianism Scale (I-CS).[62] Only the relative values derived from the H-ES were statistically or meaningfully significant in the CCP's acquaintance rape study[63] and so this article will not focus on the I-CS.

A review of the definitions reveals a potential problem, though. The definition for "egalitarian," which is supposed to be a pole of the "grid" construct, includes part of the definition of the "group" construct. Again, according to the CCP, an egalitarian is someone who "abhors social

---

available in the online appendix.

[54] *More Statistics, Less Persuasion*, *supra* note 48, at 1297.

[55] Dan M. Kahan & Donald Braman, *The Self-Defensive Cognition of Self-Defense*, 45 AM. CRIM. L. REV. 1, 28 (2008) [hereinafter *Self-Defensive Cognition*].

[56] *More Statistics, Less Persuasion*, *supra* note 48, at 1297.

[57] *Culture, Cognition, and Consent*, *supra* note 1, at 770.

[58] *More Statistics, Less Persuasion*, *supra* note 48, at 1297.

[59] *Self-Defensive Cognition*, *supra* note 55, at 28.

[60] *More Statistics, Less Persuasion*, *supra* note 48, at 1303.

[61] *Cultural Cognition and Public Policy*, *supra* note 23, at 153.

[62] *More Statistics, Less Persuasion*, *supra* note 48, at 1302-03.

[63] *Culture, Cognition, and Consent*, *supra* note 1, at 763, 779 tbl.1.

stratification, distrusts the social and political authority structures that rest on such differentiation, *and favors collective action to equalize wealth, status, and power*." That last phrase deals with collective action or assistance, and that is a part of the "group" construct. To the extent that the CCP developed scale items to measure "egalitarian," it may have inadvertently targeted the wrong construct.

C. The Hierarchy-Egalitarianism Scale Items

The next step is to evaluate whether the H-ES scale items measure the targeted construct and whether the scale items are reliable. The CCP, likely facing resource constraints and looking for a creative way to test their theories, first turned to existing data from the General Social Survey (GSS) to create a provision scale.[64] For the first study, the CCP hypothesized that a person's cultural worldviews could predict the person's risk attitude toward private gun ownership.[65]

To generate items that would measure the grid construct, the CCP looked for items in the GSS that tested attitudes about race, sexual orientation, the military, and capital punishment.[66] To generate items that would measure the group construct, the CCP looked for items that tested attitudes about public spending for regulatory and social welfare programs.[67] The CCP then used these scales to test its hypothesis that the measured latent variables would predict attitudes toward gun regulation,[68] and found that their measures did predict those attitudes.[69]

This particular study received some criticism related to its construct validity because the items did not measure constructs that were consistent with the earlier typology.[70] In response, the CCP stated that, "We are currently developing more refined measures of cultural orientation for use in [future] surveys"[71] and recognizing the difficulty in finding items that would provide content validity, stated, "We're grateful to [our critics] for focusing

---

[64] *More Statistics, Less Persuasion*, *supra* note 48, at 1302.

[65] *Id.* at 1296.

[66] *Id.* at 1303. The scale items are located in the appendix to that article. The CCP did not use any of these items in their final scale.

[67] *Id.* at 1303.

[68] *Id.* at 1302.

[69] *Id.* at 1307.

[70] *Caught in the Crossfire*, *supra* note 53, at 1407-10.

[71] *Id.* at 1408 n.46. The CCP further responded to this concern in John Gastil et al., *The 'Wildavsky Heuristic': The Cultural Orientation of Mass Public Opinion* 28-30 (Yale Law Sch. Pub. Law & Legal Theory Research Paper Series, Research Paper No. 107, Oct. 15, 2005), *available at* http://ssrn.com/abstract=834264 [hereinafter, *The 'Wildavsky Heuristic'*].

our attention on the problem. We'd be even more grateful where [they able] to figure out a way to solve it."[72]

Soon after, the CCP ran one experiment that introduced the H-ES and I-ES. This experiment resulted in two unpublished studies, *Gender, Race, and Risk Perception*[73] and *The 'Wildavsky Heuristic*'.[74] The CCP generated items for these worldview scales by looking at previously used scale items and by creating new ones. In *Gender, Race, and Risk Perception*, the CCP tells us, "Item development consisted of the adaptation of items used in previous studies . . . as well as the creation of new items".[75] In *The 'Wildavsky Heuristic'*, the CCP cite additional scales as possible sources of scale items, particularly from items in similar grid-group models.[76]

Many of these items come from scales that have their own scale development and validity problems.[77] Recognizing these issues, the CCP states that "an effort was made to create a new set of items that would exhibit better scale reliability and effectively distinguish the different dimensions of cultural orientation in factor analysis."[78]

For these first two studies, the CCP chose a four-point Likert scale but later adopted a six-point scale. The materials do not indicate that the CCP submitted this item pool to a panel of experts or included validation items, like items from a social desirability scale. The CCP does appear to have administered an item pool to a development sample, stating that they used "focus-group discussions and survey pretesting"[79] and stating that they did "extensive pretesting."[80] However, the CCP did not indicate what items were added or dropped during this process.

The CCP chose thirteen items:[81]

---

[72] *Caught in the Crossfire*, *supra* note 53, at 1411.

[73] *Gender, Race, and Risk Perception*, *supra* note 47.

[74] *The 'Wildavsky Heuristic'*, *supra* note 71.

[75] *Gender, Race, and Risk Perception*, *supra* note 47, at 12.

[76] *The 'Wildavsky Heuristic'*, *supra* note 71, at 28-31.

[77] HANDBOOK OF RISK THEORY, *supra* note 42, at 729. A reproduced item pool, along with a brief discussion of criticisms of the previous scales, and items from similar scales which were not referenced by the CCP, in available in the online appendix. *See generally* Suzanne Rippl, *Cultural Theory and Risk Perception: A Proposal for a Better Measurement*, 5 J. RISK RES. 147, 154 (2002).

[78] *The 'Wildavsky Heuristic'*, *supra* note 71 at 17. The CJP was aware of the criticisms of Karl Dake's scales. HANDBOOK OF RISK THEORY, *supra* note 42, at 730.

[79] *Gender, Race, and Risk Perception*, *supra* note 47, at 12.

[80] *The 'Wildavsky Heuristic'*, *supra* note 71, at 17. The CCP indicates that it conducted pre-screening surveys with these scales in other studies, although these prescreenings appear to have been used to measure worldviews and not to develop the scale. Dan M. Kahan et al., *Whose Eyes Are You Going to Believe? Scott v. Harris and the Perils of Cognitive Illiberalism*, 122 HARV. L. REV. 837, 859 (2009) [hereinafter *Whose Eyes*].

[81] The CCP selected at least two items verbatim from two previous scales and adopted one with slight modification. The CCP appears to have written the remaining ten items. The

Table 1: H-ES Items

Code	*Item*
HEQUAL*	We have gone too far in pushing equal rights in this country.
HREVDIS1	Nowadays it seems like there is just as much discrimination against whites as there is against blacks.
EWEALTH*	Our society would be better off if the distribution of wealth was more equal.
ERADEQ*	We need to dramatically reduce inequalities between the rich and the poor, whites and people of color, and men and women.
EDISCRIM*	Discrimination against minorities is still a very serious problem in our society.
HREVDIS2*	It seems like blacks, women, homosexuals and other groups don't want equal rights, they want special rights just for them.
HCHEATS	It seems like the criminals and welfare cheats get all the breaks, while the average citizen picks up the tab.
EDIVERS	It's old-fashioned and wrong to think that one culture's set of values is better than any other culture's way of seeing the world.
HWMNRTS	The women's rights movement has gone too far.
ESEXIST	We live in a sexist society that that is fundamentally set up to discriminate against women.
HTRADFAM	A lot of problems in our society today come from the decline in the traditional family, where the man works and the woman stays home.
HFEMININ*	Society as a whole has become too soft and feminine.
EROUGH	Parents should encourage young boys to be more sensitive and less rough and tough.

Note: * items included in short-form version of scale.

Looking first at content validity, once the construct is clearly defined, the items should cover the full domain of that construct: "In theory, a scale has content validity when its items are a randomly chosen subset of the universe of appropriate items."[82] Again, the CCP definition of "grid" is an attitude that favors or disfavors deference to traditional forms of social and political authority and is protective of the roles and status claims they entail, and where those roles and status claims are assigned on the basis of conspicuous and largely fixed attributes, such as gender, race, lineage, class, and the like.

scale found in *Gender, Race, and Risk Perception* contained the thirteen items found in the current scale plus one additional item: "EGAYMAR – A gay or lesbian couple should have just as much right to marry as any other couple." *Gender, Race, and Risk Perception*, *supra* note 47, at app. For *The Wildavsky Heuristic*, the CCP used the EGAYMAR item from the scale as a dependent variable and dropped it as an item in the scale. *The 'Wildavsky Heuristic', supra* note 71, at app. This item (EGAYMAR) was not re-introduced to the scale in later studies conducted by the CCP.

[82] DEVELLIS, *supra* note 13, at 60. *See also* Haynes et al., *supra* note 20.

Here, no items measure lineage.

Next, scale items should be representative or proportional to the facets[83] of the construct.[84] Because "[t]he items in an assessment instrument should be distributed, or weighted, in a way that reflects the relative importance of the various facets of the targeted construct",[85] the H-ES should have a distribution of items (or later, weighting during factor scoring) that reflects each facet's importance to the construct of "grid."

The distribution of items is not proportional, though. The H-ES has thirteen items that cover four of those facets: gender, race, sexual orientation, and class. Five of the thirteen scale items solely measure attitudes about gender roles (HWMNRTS, ESEXIST, HTRADFAM, HFEMININ, EROUGH) and two other items (ERADEQ, HREVDIS2) include gender among other dimensions – for a total of more than half of the items. One item (EDISCRIM) solely measures race, while two other items include a race dimension (ERADEQ, HREVDIS2), for a total of three. Two items appear to have been written to measure class (EWEALTH, HCHEATS) and class is included as a dimension in one other item (ERADEQ), for a total of three. One item (HREVDIS2) contains sexual orientation, among other facets.

Further, a single facet should not have overly redundant items.[86] If the aggregate score for a scale is disproportionately influenced by any one facet of the construct, the scale will lack content validity.[87] When a scale is designed to measure an entire global construct but has several items that measure a specific subconstruct, the result may be "to pull the item set as a whole away from the intended latent variable . . . to an alternative, more specific latent variable".[88]

This scale has redundancy problems. In particular, the gender items stand out. With up to seven of the thirteen items drawing on gender hierarchy, the H-ES may really be measuring that subconstruct.

Continuing on with content validity, some of the items may not be measuring the intended construct. The "grid" items should only be measuring the respondent's agreement with social structure (a social order that features differentiation and stratification of social roles based on fixed characteristics) and should not be measuring agreement with collective interference or collective action. However, some items may also be tapping into the "group" dimension, likely because the definition of "egalitarian" included a "group"

---

[83] A facet is a "dimension of interest [that] is a potential source of variation." DEVELLIS, *supra* note 13, at 56.

[84] Haynes et al., *supra* note 20, at 239; Clark & Watson, *supra* note 13, at 311.

[85] Haynes et al., *supra* note 20, at 245; *see also* Clark & Watson, *supra* note 13, at 311.

[86] DEVELLIS, *supra* note 13, at 79.

[87] Haynes et al., *supra* note 20, at 240.

[88] DEVELLIS, *supra* note 13, at 79.

facet. Some items appear to tap into "group" fairly directly (EWEALTH, ERADEQ) while others (HEQUAL, HCHEATS, ERADEQ, HREVDIS2, ESEXIST, EROUGH, HWMNRTS) use words like "we need to" or "parents should" that seem to indicate agreement with collective action, collective interference, group identity, or group movements.

Turning to item reliability, two of the items (ERADEQ, HREVDIS2) are complex or "double-barreled."[89] Complex items are ones that "convey two or more ideas so that endorsement of the item might refer to either or both ideas."[90] Complex items have reliability problems: "respondents will interpret complex items in different ways; accordingly, their responses will reflect the heterogeneity of their interpretations."[91] It is difficult to know whether people with the same belief systems will answer these two items in the same way, and it is difficult to understand what the items actually measure. For example, the item language for ERADEQ is, "We need to dramatically reduce inequalities between the rich and the poor, whites and people of color, and men and women." Some people might believe in race or income equality but still believe in traditional gender roles, or some may believe in gender equality (because they have more exposure to people of the opposite sex) but still have latent racist beliefs (because of a lack of exposure to other races).[92] A response to this item of "slightly agree" could mean that the respondent believes in race equality but not gender equality, or it could mean that the respondent believes in gender equality but not race equality. Further, the next respondent with the same belief systems as the last respondent might answer the item in a different way.

Other items are ambiguous. In general, "a good item should be unambiguous."[93] For example, the item language for HCHEATS is, "It seems like the criminals and welfare cheats get all the breaks, while the average citizen picks up the tab." This item appears to have been written to test class hierarchy, but it may actually test race hierarchy if many respondents assume that criminals and welfare cheats are minorities.[94] Likewise, the item language for EDIVERS is, "It's old-fashioned and wrong

---

[89] *Id.* at 81-82; Clark & Watson, *supra* note 13, at 312. This is also called construct-irrelevant difficulty. Samuel Messick, *Validity of Psychological Assessment: Validation of Inferences From Persons' Responses and Performances as Scientific Inquiry Into Score Meaning*, 50 AM. PSYCHOL. 741, 742 (1995).

[90] DEVELLIS, *supra* note 13, at 82.

[91] Clark & Watson, *supra* note 13, at 312.

[92] Racist and sexist beliefs are correlated and have similarities, but also have differences in their subconstructs. *See generally* Janet K. Swim et al., *Sexism and Racism: Old-Fashioned and Modern Prejudices*, 68 J. PERSONALITY & SOC. PSYCHOL. 199 (1995).

[93] DEVELLIS, *supra* note 13, at 81.

[94] Exploratory factor analysis suggests that this item measures race hierarchy. *See* discussion *infra* Part V.A.

to think that one culture's set of values is better than any other culture's way of seeing the world." The word "culture" can mean many things to many people, and this item could be measuring beliefs about race, or national origin, or lineage.

Also, two of the items (HEQUAL, EDIVERS) are abstract or global and do not directly measure any facet, while other items list specific facets. Designing scales to measure abstract, global, or general constructs is perfectly acceptable: "Scales can be developed to assess constructs at each of many levels of abstraction."[95] But, the "scale's content should reflect the conceptual definition application *to that scale*"[96] and scale developers should "select item wordings that correspond[] with the intended level of variable specificity."[97]

Mixing abstract and specific items can cause problems. Subclusters can form on the abstract items based on their nonspecificity.[98] The abstract items can also become ambiguous when surrounded by specific facets. For example, with the item HEQUAL ("We have gone too far in pushing equal rights in this country"), a respondent might agree or disagree with that item because of the influence of a particular facet that has already been listed in the surrounding items on that scale. The respondent might think that we have not gone far enough with race equality but agree with this item because he believes the women's rights movement has gone too far, and the surrounding items have been heavily weighted toward gender roles and the respondent now has gender roles on the mind.

To sum, the scale appears to have content validity issues[99] because the entire content domain is not represented; the items have proportionality and redundancy issues; and some items may measure a different construct. Because of the redundancy problem, it may turn out that the gender items have narrowed the construct that the scale is measuring from global hierarchy to gender hierarchy. Further, several of the items have reliability issues because they are ambiguous, complex, or shift between specific and abstract forms.

D. The CCP's Item Evaluation

For item evaluation, the CCP initially reported an alpha of .80 for the H-

---

[95] Clark & Watson, *supra* note 13, at 310.

[96] DEVELLIS, *supra* note 13, at 60 (emphasis in original).

[97] *Id.* at 115.

[98] *Id.* at 80.

[99] *See also* Eero Olli, *Rejected Cultural Biases Shape Our Political Views* 290 (Mar. 29, 2012) (unpublished Ph.D. dissertation, University of Bergen) (on file with author).

ES items[100] and has consistently found a high coefficient alpha for the H-ES.[101] (In *Culture, Cognition, and Consent*, Kahan reports an alpha of .89).[102] In *The 'Wildavsky Heuristic'*, the CCP indicated that it might have conducted factor analysis but does not report the results of that analysis.[103] The CCP also conducted structural equation modeling in one study, but the modeling was not related to item evaluation.[104] These discussions are the only reported discussions on the full H-ES item evaluation.[105]

Scale developers often use coefficient alpha to test for reliability: "Alpha is defined as the proportion of a scale's total variation that is attributable to the common source, presumably the true score of the latent variable underlying the items."[106] If a scale measures only one latent variable and the scale items are highly correlated with each other, then we attribute that correlation to the latent variable.

As a measure of reliability, however, coefficient alpha has a fundamental assumption: the scale must be unidimensional.[107] As discussed above, the H-ES may not be. Rather, the scale may be multi-dimensional, with gender, class, and race hierarchy all forming facets, and the scale may also measure

---

[100] *The 'Wildavsky Heuristic'*, *supra* note 71 at 17.

[101] For the fourteen-item scale, the CCP reported a coefficient alpha for H-E of .81. *Gender, Race, and Risk Perception*, *supra* note 47, at tbl.1. Later studies using the thirteen-item scale also reported coefficient alpha for the H-ES: *Affect, Values, and Nanotechnology Risk Perceptions*, supra note 51, at 11 (.81); Dan M. Kahan et al., *Cultural Cognition of the Risks and Benefits of Nanotechnology*, 4 NATURE NANOTECHNOLOGY 87, 89 (2008) (.81) [hereinafter *Risks and Benefits of Nanotechnology*]; *Self-Defensive Cognition*, *supra* note 55, at 28 (.82); *Whose Eyes*, *supra* note 80, at 860 (.85); Dan M. Kahan et al., *Risk and Culture: Is Synthetic Biology Different?* 9 (Cultural Cognition Project Working Paper, Paper No. 29 (2009)) (.86) [hereinafter, *Is Synthetic Biology Different?*]; Dan M. Kahan et al., *Cultural Cognition and Public Policy: the Case of Outpatient Commitment Laws*, 34 LAW. HUM. BEHAV. 118, 125 (2010) (.89) [hereinafter, *Outpatient Commitment Laws*]; Dan M. Kahan et al., *Who Fears the HPV Vaccine, Who Doesn't, and Why? An Experimental Study on the Mechanisms of Cultural Cognition*, 34 LAW. HUM. BEHAV. 501, 505 (2010) (.80) [hereinafter, *Who Fears the HPV Vaccine*].

[102] *Culture, Cognition, and Consent*, *supra* note 1, at 770.

[103] *The 'Wildavsky Heuristic'*, *supra* note 71 at 17. In a later article that summarized that article, the CCP states that "Because the cultural orientations were not conceptualized as uncorrelated, mean item scores (rather than factor scores) were used to generate reliable measures of egalitarian-hierarchy (alpha = .82) and individualism-solidarism (alpha = .79)." John Gastil et al., *The Cultural Orientation of Mass Political Opinion*, 44 POL. SCI. & POL. 711, 712 (2011).

[104] *Self-Defensive Cognition*, *supra* note 55.

[105] The CCP later modified this scale to form a short-form scale and discuss the reliability and validity of that modified scale. *See* discussion *infra* note 116.

[106] DEVELLIS, *supra* note 13, at 37.

[107] Smith & McCarthy, *supra* note 107, at 302 ("For a scale to measure a unidimensional construct, its items must be parallel, alternative indicators of the same, underlying construct.")

some part of the "group" construct, too.

A scale can capture more than one dimension (say, racial, gender, and class attitudes) and still return a high coefficient alpha.[108] While the CCP reports high coefficient alpha values, a "high coefficient alpha does not indicate unidimensionality"[109] and "a relatively high alpha is no guarantee that all the items reflect the influence of a single latent variable."[110] Research has shown that "acceptable alpha levels can be obtained by aggregating distinct but correlated subscales".[111] Further,

> [P]sychometricians long have disavowed the practice of using reliability indices to establish the homogeneity of a scale. To understand why this is so, it is necessary to distinguish between internal consistency on the one hand and homogeneity or unidimensionality on the other. *Internal consistency* refers to the overall degree to which the items that make up a scale are intercorrelated, whereas *homogeneity* and *unidimensionality* indicate whether the scale items assess a single underlying factor or construct. Internal consistency is a necessary but not sufficient condition for homogeneity or unidimensionality. In other words, a scale cannot be homogeneous unless all of its items are interrelated, but . . . a scale can contain many interrelated items and still not be unidimensional.[112]

The gender item redundancy issue within the H-ES can also affect unidimensionality. When a scale has been designed to measure a general construct but it has several items related to a subconstruct, that can mean that "[c]orrelations among those items are likely to be greater than correlations between those items and others not related [to the subconstruct]. This can . . . undermine the unidimensionality of the item set".[113]

Failure to "identify the measure's hierarchical or aggregational structure . . . could lead to inaccurate specifications of theory as well as misleading correlational and experimental findings."[114] A typical failure in many

---

[108] John Hattie, *Methodology Review: Assessing Unidimensionality of Tests and Items*, 9 APPLIED PSYCHOL. MEASUREMENT 139, 144 (1985); DEVELLIS, *supra* note 13, at 116.

[109] Smith & McCarthy, *supra* note 107, at 303.

[110] DEVELLIS, *supra* note 13, at 116.

[111] Smith & McCarthy, *supra* note 107, at 301; *id*. at 303 ("Because coefficient alpha is influenced by both internal consistency and scale length, it can be high when two internally consistent subscales, themselves only modestly interrelated, are combined").

[112] Clark & Watson, *supra* note 13, at 315 (emphasis in original).

[113] DEVELLIS, *supra* note 13, at 79.

[114] Smith & McCarthy, *supra* note 107, at 300.

manuscripts submitted to one journal "was one in which investigators reported a coefficient alpha and apparently presumed the value indicated unidimensionality without testing that assumption".[115]

That appears to have happened here.[116] The CCP has not published the scale's psychometrics so we do not know if the scale is unidimensional or multidimensional. Determining whether a scale measures a global construct or instead returns an aggregate of related but distinct factors is usually done with factor analysis.[117] The exploratory factor analysis in Section V will help expose the scale's factor structure, particularly whether it is unidimensional or multidimensional.

IV. KAHAN'S VALIDITY ARGUMENT IN *CULTURE, COGNITION, AND CONSENT*

Within the sexual assault context, Kahan recognized that its "grid" variable differed from the gender role constructs used by other researchers

---

[115] Smith & McCarthy, *supra* note 107, at 301.

[116] The CCP ran some studies of a short-form version of the scale, using six items from the H-ES and six items form the I-C Scale. *Cultural Cognition of Scientific Consensus*, *supra* note 51, at 151 and app.1; Dan M. Kahan et al., *"They Saw a Protest": Cognitive Illiberalism and the Speech-Conduct Distinction*, 64 STAN. L. REV. 851, 869 (2012) [hereinafter, *"They Saw a Protest"*]; HANDBOOK OF RISK THEORY, *supra* note 42. However, these results do not resolve these issues. The CCP reported an alpha for the short version of the H-E Scale of .87 and stated that the items "loaded appropriately on two separate factors, which were used as predictors for the study." *Cultural Cognition of Scientific Consensus*, *supra* note 51, at 151; *see also "They Saw a Protest"*, *supra*, at 869-70; HANDBOOK OF RISK THEORY, *supra* note 42. However, on the orthogonally rotated factor matrix, the items actually plotted on four separate factors, not two, with the H and E items plotting separately. HANDBOOK OF RISK THEORY, *supra* note 42, at 730 fig. 28.5. The CCP probably should have used oblique rotation because the factors are correlated. Leandre R. Fabrigar et al., *Evaluating the Use of Exploratory Factor Analysis in Psychological Research*, 5 PSYCHOL. METHODS 272, 281 (1999)). In a recent study, the CCP reduced the H-E scale to just two items (HEQUAL and EWEALTH) and reported that factor analysis showed that they loaded on the same factor with a coefficient alpha of 0.73. Dan M. Kahan et al., *"Ideology" or "Situational Sense"? An Experimental Investigation of Motivated Reasoning and Professional Judgment* 20 (Cultural Cognition Project Working Paper No. 63, 2015).

In general, scale developers should run complete scale development on short-form scales: "We advise against [reducing scales to short forms]: the psychometric properties of a measure cannot be imputed to a short form without empirical testing. Often, use of abbreviated measures attenuates reliability." Smith & McCarthy, *supra* note 107, at 306. Reducing scales to short-forms also can cause issue with content validity: "Even more frequently, internal consistency is preserved but validity is attenuated because of reduced coverage of the target construct." *Id.* Global constructs have larger universes and covering the full domain can be difficult. Not covering the entire domain leads to construct underrepresentation. Messick, *supra* note 89, at 742.

[117] Smith & McCarthy, *supra* note 107, at 303 ("there are direct means of assessing the degree of subscale covariance (e.g., examination of correlation matrices and use of confirmatory factor analysis to test the degree of loss of model fit when combining scales.")

who study sexual assault and offered a lengthy reconciliation. Kahan stated:

> Hierarchy is comparable but not identical to the attitudinal measures used to characterize subjects' gender-norm attitudes in [other studies of sexual assault]. Like those measures, Hierarchy includes items that relate to traditional gender roles and sexual equality . . . However, it also includes items that relate to other dimensions of social stratification unrelated to the gender-norm measures used in those studies . . .[118]

Kahan argued that this construct was valid because of the high coefficient alpha it found for these scale items: "Because the reliability of Hierarchy as a latent attitudinal measure indicates a high degree of affinity between hierarchical gender attitudes and hierarchical attitudes generally, there is no conceptual difficulty in using Hierarchy to test hypotheses related to the former."[119] However, the discussion above about coefficient alpha does not support this statement.

Kahan then argued that its global construct would better explain the variance than just a gender role facet: "Indeed, positive results obtained by the use of Hierarchy are arguably stronger than ones based on gender-role attitudinal scales. Hierarchy measures a disposition more general than those measured by gender-role scales and is conceptually more remote from the study's dependent variables, which themselves relate to perceptions of sexual behavior."[120]

Kahan cites Paul Slovic and Ellen Peters for the proposition that "the influence of distal [or global] variables is ordinarily smaller but more important than the influence of proximal variables".[121] Those authors were themselves responding to criticism about their choice of using a global variable, but there, the authors were careful to point out that they "purposely selected worldview items to minimize semantic overlap with the risk attitudes and perception being explained."[122] They chose global variables because they wanted to prevent their independent variables from looking too much like their dependent variables, which is a different issue than Kahan faced in his study.

Choosing to investigate a global variable versus a specific variable (called

---

[118] *Culture, Cognition, and Consent*, *supra* note 1, at 770.

[119] *Id.*

[120] *Id.*

[121] *Id.* at 770 n.157, *citing* Paul Slovic & Ellen Peters, *The Importance of Worldviews in Risk Perception*, 3 Risk Decision & Pol'y 165, 168 (1998).

[122] Slovic & Ellen, *supra* note 121, at 169.

construct hierarchy)[123] is not itself a problem, and specific variables are not better or worse that global variables.[124] Further, specific facets can be combined and analyzed at an aggregate, global level. But, researchers can do that only if the correlations between the facets and the dependent variable do not differ from the main effect of the aggregated score.[125] If the facets perform differently than the aggregated score, the researcher needs to report that result.[126]

The structural equation modeling in Section VI will explore whether the scale has different facets and whether those facets perform differently than the hypothesized global variable.

V. EXPLORATORY FACTOR ANALYSIS OF THE HIERARCHY-EGALITARIANISM SCALE

I conducted exploratory factor analysis (EFA) to better understand how the items related to each other and the scale's possible factor structure, which would then inform decisions about the theoretical model for the structural equation modeling (SEM).

A. Survey Administration and Data Screening

The data comes from a survey that was administered in 2009. Kahan used a private firm to administer the survey to 1,500 people in the United States. The survey was conducted online using a pool of over one million people who are paid to participate in these surveys. The firm used a demographic-matching methodology that ensured that the sample was representative of the general population so weighting is not necessary.[127]

Using Excel, I screened the data to see if any observations were missing data over 10%. I screened for unengaged respondents by running the

---

[123] Construct hierarchy refers to the generality or specificity of the construct that the scale developer is trying to measure. *See* Andrew L. Comrey, *Factor-Analytic Methods of Scale Development in Personality and Clinical Psychology*, 56 J. CONSULTING & CLINICAL PSYCHOL. 754, 755-56 (1988); Clark & Watson, *supra* note 13, at 310; Gregory T. Smith & Dennis M. McCarthy, *Methodological Considerations in the Refinement of Clinical Assessment Instruments*, 7 PSYCHOL. ASSESSMENT 300, 301 (1995). For our purposes, this is an unfortunate term because the CCP sometimes calls the construct that we are exploring "hierarchy." I will instead use the terms "global construct" and "specific construct" to capture the concept rather than using "construct hierarchy." *See also* DEVELLIS, *supra* note 13, at 79.

[124] The level of specificity of the scale should generally match the research question, though. DeVellis, *supra* note 13, at 75.

[125] Smith & McCarthy, *supra* note 107, at 303.

[126] *Id.*

[127] *Culture, Cognition, and Consent*, *supra* note 1, at 765.

standard deviation for each respondent's data and looking closely at those with low standard deviations. I removed those that I found. I screened the variables for outliers and found two in the variable EROUGH. I deleted those data points but retained the observations. The variable "PID7" (7-point Likert scale that measured political identification from liberal through conservative) was missing data at 4%. No other variable had missing data over 1%. After screening, n = 1487. I randomly split the sample in half so that one half could be used for exploratory factor analysis (n = 770) and the other half for structural equation modeling (n = 717).[128]

B. Exploratory Factor Analysis

I used SPSS v. 21. I excluded missing data pairwise. The data is not normally distributed[129] and normality is a strict assumption for maximum likelihood factor solutions,[130] which is the common way of conducting exploratory factor analysis. Therefore, I used principal factor extraction method (in SPSS, principal axis factors)[131] because principal factor methods do not have a distribution assumption.[132] I expected that any factors that were extracted would be correlated so I chose an oblique rotation (Promax). I did not reverse code the "E" items so that it would be easier to spot criterion validity.

The correlation matrix (provided in Table 2) contained many items with correlations of r = .3 or greater.[133]

---

[128] Fabrigar et al., *supra* note 116, at 277.

[129] I conducted a visual inspection of the histograms for all of the H-E items. Only ESEXIST and EROUGH looked somewhat normal. A Shapiro-Wilk's test ($p > .05$) showed that all of the variables were non-normal (all had p values $< .001$). Many of the items were severely skewed (z-value > 2). All items except for HCHEATS showed negative kurtosis with many showing severe negative kurtosis (z-value > 7), Fabrigar et al., *supra*, at 283, meaning the distribution departed from a bell curve toward a uniform or flat distribution. Normality values are available in the online appendix.

[130] Frank J. Floyd & Keith F. Widaman, *Factor Analysis in the Development and Refinement of Clinical Assessment Instruments*, 7 PSCYHOL. ASSESSMENT 286, 289 (1995).

[131] *Id.*; TIMOTHY A. BROWN, CONFIRMATORY FACTOR ANALYSIS FOR APPLIED RESEARCH 379, 387 (2006); Anna B. Costello & Jason W. Osborne, *Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis*, 10 PRACTICAL ASSESSMENT, RES. & EVALUATION 1, 2 (2005). I also ran exploratory factor analysis using maximum likelihood and unweighted mean squares. The results were substantially the same: "[W]hen the common factor model holds reasonably well in the population and severe violations of distributional assumptions are not present, solutions provided by these methods are usually very similar." Fabrigar et al., *supra* note 116, at 277.

[132] Fabrigar et al., *supra* note 116, at 277.

[133] The Kaiser-Meyer-Olkin Measure of Sampling Adequacy was above 0.6 (0.92), indicating a high degree of common variance. Bartlett's Test of Sphericity showed

Table 2: Correlation Matrix for H-ES Exploratory Factor Analysis

Variable	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
1. HEQUAL	-												
2. HREVDIS1	.56	-											
3. EWEALTH	-.40	-.27	-										
4. ERADEQ	-.43	-.31	.69	-									
5. EDISCRIM	-.48	-.41	.44	.57	-								
6. HREVDIS2	.62	.59	-.41	-.44	-.51	-							
7. HCHEATS	.46	.50	-.24	-.26	-.30	.56	-						
8. EDIVERS	-.26	-.13	.35	.40	.29	-.29	-.12	-					
9. HWMNRTS	.57	.44	-.41	-.44	-.44	.61	.41	-.32	-				
10. ESEXIST	-.30	-.23	.41	.46	.46	-.41	-.20	.36	-.42	-			
11. HTRADFAM	.40	.36	-.29	-.33	-.32	.49	.39	-.27	.50	-.24	-		
12. HFEMININ	.44	.40	-.36	-.40	-.40	.52	.39	-.30	.54	-.40	.47	-	
13. EROUGH	-.20	-.18	.40	.38	.37	-.29	-.14	.27	-.28	.40	-.23	-.38	-

Two factors were extracted. Both of the eigenvalues for the two factors exceeded the value set by parallel analysis[134] and the scree plot was consistent with two factors. All of the items loaded on two factors, with all of the "H" items loading on factor 1 and all of the "E" items loading on factor 2. All of the items loaded above .4.[135] The pattern matrix and communalities table are provided in Table 3.

Table 3: Exploratory Factor Analysis Pattern Matrix and Communalities

	Pattern Matrix Factor		*Communalities*	
Variable	*1*	*2*	*Initial*	*Extraction*
HREVDIS1	.80		.45	.51
HCHEATS	.79		.39	.46
HREVDIS2	.79		.62	.71
HEQUAL	.70		.52	.55
HWMNRTS	.56		.52	.54
HTRADFAM	.53		.36	.35
HFEMININ	.45		.44	.45
ERADEQ		.83	.59	.64
EWEALTH		.76	.52	.53

significance ($p < .001$), meaning the null hypothesis that the correlation matrix is an identity matrix (where the variables are noncollinear) and that any non-zero correlations are due to sampling error is rejected. These indices support the factorability of the correlation matrix.

[134] I ran Monte Carlo PCA for Parallel Analysis, with variables set at 13, number of subjects at 750, and 1000 iterations. The criterion eigenvalue for the first extracted factor was 1.22, SD = 0.03 and the actual eigenvalue was 5.70; the criterion for the second was 1.17, SD = 0.02 and the actual was 1.50; the criterion for the third was 1.13, SD = .02 and the actual was 0.85.

[135] Floyd & Widaman, *supra* note 130, at 294 (items should load greater than .3 or .4); J. Raubenheimer, *An Item Selection Procedure to Maximise Scale Reliability and Validity*, 54 J. INDUS. PSYCHOL. 59, 61 (2004) (items should load greater than .4.)

ESEXIST	.63	.37	.41
EROUGH	.59	.28	.30
EDIVERS	.54	.24	.26
EDISCRIM	.51	.47	.47

The factors were strongly, negatively correlated (r = -.66). Coefficient alpha for factor 1 is .87 and for factor 2 is .82. Two items, EDIVERS and EROUGH, had low communalities which indicates that they do not relate well to the other items or are unreliable.[136] The model explained less than 50% of variance (48%)[137] which indicates that there is a lot of unique variance and random error, possibly because of the issues related to the scale validity and reliability that are discussed above.[138]

The EFA provides a two-factor solution, but that two-factor solution does not suggest that the factors are related to particular facets of "grid" like race, gender, or class. That might be because there are not enough items in the scale to measure each subfactor: "[F]actor analysis might establish that the items can be subdivided into several subscales but that the initial pool does not contain enough items to assess each of these content domains reliably."[139] Generally, to detect facets or individual factors, researchers need three to six items per facet or factor,[140] and here, we only have one item in the scale that solely measures race and two that are designed to solely measures class. Sexual orientation is mentioned in only one item and there are no items on lineage.

Importantly, the exploratory factor analysis does not support the CCP's unidimensional scale design. If the H-ES was unidimensional, all of the items should load on one factor, with "grid" serving as the latent variable. The

---

[136] Fabrigar et al., *supra* note 116, at 274: "There are a number of reasons why communalities for measured variables might be low. One obvious reason is low reliability. As explained later, variance due to random error cannot, by definition, be explained by common factors. Because of this, variables with low reliability will have low communalities and thus should be avoided."

[137] Floyd & Widaman, *supra* note 130, at 295 (minimum of .50 and recommends above .80).

[138] *Id.* ("[A]ccounting for relatively little variance challenges the relative importance of common factors as opposed to the specific factor variance associated with individual variables.")

[139] Clark & Watson, *supra* note 13, at 311.

[140] Floyd & Widaman, *supra* note 130, at 292 ("In general, three variables per factor are needed to identify common factors"); Fabrigar et al., *supra* note 116, at 273 ("Research suggests that [exploratory factor analysis] procedures provide more accurate results when each common factor is represented by multiple measured variables in the analysis . . . Methodologists have recommended that at least three to five measured variables representing each common factor be included in each study"). Fabrigar and colleagues recommend four to six. *Id*. at 282.

loading reported by the EFA was not due to reverse coding (I ran the EFA both with and without reverse coding and the results were the same) or because the items were written in opposite directions (for example, some items but not others using "not" language),[141] both of which can cause a single dimension to load on two separate poles. Rather, the EFA suggests that the H-ES might roughly measure two separate constructs.

VI. STRUCTURAL EQUATION MODELING

The exploratory factor analysis suggests that the items do not measure a unidimensional construct and that there may be reliability issues with the items. I used structural equation modeling to see if the scale items could still deliver useful information.

Structural equation modeling has two components: a measurement model and a structural model.[142] The measurement model is usually a theory-driven confirmatory factor analysis, which is contrasted to the data-driven exploratory factor analysis that I ran earlier.[143] The structural model "displays the interrelations among latent constructs and observed variables in the proposed model as a succession of structural equations – akin to running several regression equations."[144] Statistical software then returns various model fit indices, which indicate the degree to which the specifications of the theory-driven model "is consistent with the pattern of variances and covariances from a set of observed data."[145] Most indices "reflect the improvement in fit of a specified model . . . over the independence model, in which all structural parameters are fixed at zero."[146]

I organized the items into subconstructs based on an analysis of item wording and the correlation matrix from the EFA. For the gender facet, I directed the model to load the gender items HWMNRTS, ESEXIST, HTRADFAM, FFMININ, and EROUGH onto a single factor. For the race facet, I directed EDISCRIM (purely a race item), HREVDIS1 (purely a race item), HREVDIS2 (the item starts with a race facet and is highly correlated with EDISCRIM and HREVDIS1), HEQUAL (highly correlated with EDISCRIM, HREVDIS1, HREVDIS2, and HCHEATS, although it is also highly correlated with HWMNRTS), HCHEATS (low correlation with

---

[141] Comrey, *supra* note 123, at 758.

[142] James B. Schreiber et al., *Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review*, 99 J. EDUC. RES. 323, 325 (2006).

[143] *Id.* at 323.

[144] *Id.* at 325.

[145] Rick H. Hoyle, *The Structural Equation Modeling Approach: Basic Concepts and Fundamental Issues*, *in* STRUCTURAL EQUATION MODELING: CONCEPTS, ISSUES, AND APPLICATIONS 1, 3 (Rick H. Hoyle, ed. 1995).

[146] *Id.* at 7.

EWEALTH and a strong correlation with the race items), and EDIVERS (this item had low correlations with most items and I theorized that many respondents would think of race when they think of "culture") to load on a single facet. I directed EWEALTH and ERADEQ (the item starts with an income facet and is highly correlated with EWEALTH) to load on a class facet.

A. Confirmatory Factor Analysis

I used Mplus version 6.12 statistical software. Because of multivariate nonnormality (the Mardia's coefficient was significant), I ran a bootstrap with 2000 replicates. I used FIML to deal with missing data. As noted above, n = 717.

The initial model fit was not good. The chi-square value for the overall model fit was significant (chi-square (62) = 494.72, p < .001).[147] However, for large samples sizes like the one studied here, "it can happen the chi-square test is failed even though differences between observed and predicted covariances are slight."[148] Further, chi-square "is viewed by most as overly strict given its power to detect even trivial deviations of data from the proposed model"[149] and so researchers should also look to other indices to evaluate model fit. Here, examination of other indices also showed an unacceptable model fit (RMSEA = .10, p(close) < .001; CFI = .89; TLI = .86; SRMR = .06).[150]

Modification indices indicated that ERADEQ had cross loading issues, meaning that it has substantial shared variance with all three facets (probably because it is a complex item that asks about race, gender, and class hierarchy), so I removed it. This meant that the class facet only had one observed variable (EWEALTH) so I removed this item and facet from the model. I then removed EDIVERS for cross loading issues which are probably the result of the ambiguous term "culture." I correlated the error terms for

---

[147] In SEM, the chi-square statistic is really a "badness of fit" index. Hoyle, *supra* note X, at 7. A large chi-square and a small p value (here, less than .01) means that there is a statistically significant difference between the estimated variance-covariance matrix and the actual variance-covariance matrix. Here, this statistic means that we reject the exact-fit hypothesis at the .01 level and suggests a lack of fit between the hypothesized model and the data. *See* REX B. KLINE, PRINCIPLES AND PRACTICE OF STRUCTURAL EQUATION MODELING 199-201 (3d ed. 2011).

[148] *Id.* at 201.

[149] Ralph O. Mueller & Gregory R. Hancock, *Structural Equation Modeling*, *in* THE REVIEWER'S GUIDE TO QUANTITATIVE METHODS IN THE SOCIAL SCIENCES 371, 379 (Gregory R. Hancock & Ralph O. Mueller, eds. 2010).

[150] Cutoff criteria: RMSEA < .06; CFI ≥ .95; TLI ≥ .95; SRMR ≤ .08. Schreiber et al., *supra* note 142, at 330 tbl.2.

EROUGH and ESEXIST, and EROUGH and HFEMININ because the modification indices indicated that these items had a lot of shared error variance. This indicates that the gender facet has possible subscales. The modification indices also indicated shared error variance between HCHEATS and HREVDIS1, so I correlated the error terms. This indicates that the race facet has possible subscales. I then removed EDISCRIM because of cross loading issues with two of the gender items.

The resulting model has two factors with the items loading appropriately on their respective factors (.43 - .79 for gender and .58 - .88 for race). The chi-square value for the overall model fit was significant (chi-square (23) = 48.24, $p < .001$), suggesting a lack of fit between the hypothesized model and the data. However, examination of other indices also showed an acceptable model fit (RMSEA = .04, p(close) = .87; CFI = .99; TLI = .98; SRMR = .02).[151] The gender and race constructs were highly correlated (r = .87), which indicates that they may be subconstructs of a more global construct.[152]

B. The Structural Model

I used the results from the confirmatory factor analysis to establish the structural model. In addition to the nine observed variables (the scale items) loading on two latent variables (gender and race) with one exogenous variable (GUILT), I included one control variable – political identification. The CCP reported that political identification as a liberal or conservative did not meaningfully influence the responses; however, other research has found that this variable predicts rape myth acceptance[153] so I included it as a control variable in the model.[154]

---

[151] Keith F. Widaman & Jane S. Thompson, *On Specifying the Null Model for Incremental Fit Indices in Structural Equation Modeling*, 8 PSYCHOL. METHODS 16 (2003).

[152] To contrast to with structure, I also ran confirmatory factor analysis on the structure revealed by the exploratory factor analysis, where the "H" items loaded on one factor and the "E" items loaded on another. The model fit was not good. The chi-square value for the overall model fit was significant (chi-square (64) = 355.54, $p < .001$), suggesting a lack of fit between the hypothesized model and the data. Examination of other indices also showed an unacceptable model fit (RMSEA = .08, p(close) < .001; CFI = .93; TLI = .91; SRMR = .05).

[153] *See* Anderson et al., *supra* note 11, at 312; William D. Walker et al., *Authoritarianism and Sexual Aggression*, J. PERSONALITY & SOC. PSYCHOL. 1036, 1038 (1993) (using the Right Wing Authoritarianism scale).

[154] The CCP reported that race, age, and education did not meaningfully influence responses to the guilt variable, *Culture, Cognition, and Consent*, *supra* note 1, at 782. This is consistent with other research so I did not include these as control variables. *See* Kimberly A. Lonsway & Louise F. Fitzgerald, *Rape Myths: In Review*, 18 PSYCHOL. WOMEN Q. 133, 142-45 (1994). The CCP reported that sex did not meaningfully influence responses. This is inconsistent with other research, *see* Lonsway & Fitzgerald, *supra*, but I did not include it

As noted above, Kahan applied a treatment condition by randomly dividing the subjects into five groups of 300 and giving each of them one of five legal conditions (either no law to apply, or four different versions of law to apply).[155] With the guilty variable converted to binary, Kahan reported that only one of these conditions was statistically significant.[156] I ran a univariate general linear model (ANOVA) with the treatment conditions as the independent variable and GUILTY as the dependent variable. The main effect was not significant ($F(4,709) = .40$, $p = .81$, eta squared < .01), nor were any of the pairwise comparisons. Therefore, I did not include the treatment condition variable as a control variable in the model.

The chi-square value for the overall model fit was significant (chi-square $(36) = 84.02$, $p < .001$). However, examination of other indices showed an acceptable model fit: RMSEA = .04, p(close) = .82; CFI = .98; TLI = .98; SRMR = .03. A correlation table is provided in Table 4 and the theoretical model is provided in Figure 3.

Table 4: Correlation Matrix for Structural Model

Variable	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
1. HFEMININ	-								
2. HWMNRTS	.57	-							
3. HTRADFAM	.45	.49	-						
4. ESEXIST	-.36	-.40	-.24	-					
5. EROUGH	-.47	-.33	-.19	.43	-				
6. HREVDIS1	.41	.46	.33	-.29	-.24	-			
7. HREVDIS2	.56	.60	.47	-.38	-.36	.66	-		
8. HEQUAL	.53	.60	.41	-.38	-.33	.57	.69	-	
9. HCHEATS	.36	.37	.31	-.23	-.22	.50	.52	.43	-

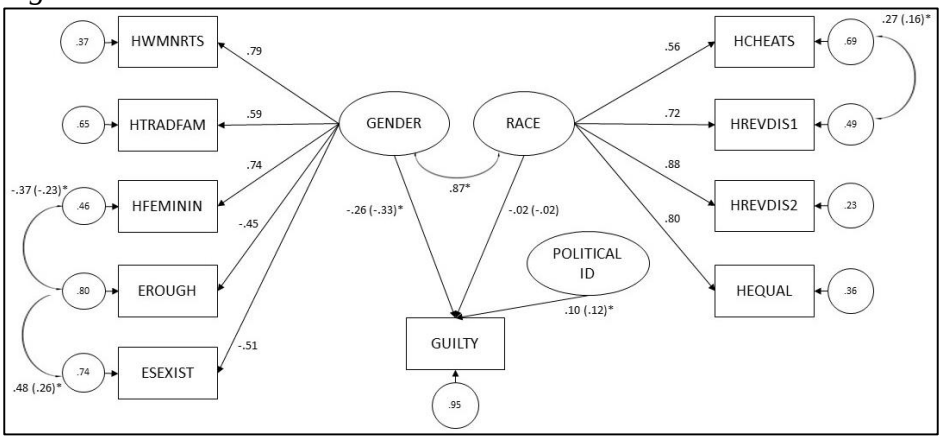
Note: Spearman's rho correlations, all significant (two-tailed) at $p < .01$.

in this model because gender does not itself involve a competing belief system to the two other latent variables in the model. After I ran the model, I later included gender and it did not improve model fit and it was a statistically insignificant predictor of the outcome judgment.

[155] *Culture, Cognition, and Consent*, *supra* note 1, at 767-68.

[156] *Id.* at 778, 779 tbl.1.

Figure 3: Theoretical model



Note: Standardized path coefficients are in parentheses. * $p < .01$.

The data suggests that the gender items account for the variance in the outcome judgment and not the race items. The gender facet was a significant (statistically and meaningfully) predictor of agreement with the outcome judgment of guilty. For every increase in one standard deviation from the mean of "Gender" (toward being more traditional), the mean of "Guilty" would be expected to decrease by .33 in its own standard deviations from its mean, while holding all other relevant connections constant. The "Race" facet was not statistically significant. The "Political ID" facet was statistically significant but had a smaller effect. For every increase in one standard deviation from the mean of "Political ID" (toward being more liberal), the mean of "Guilty" would be expected to increase by .12 in its own standard deviations from its mean, while holding all other relevant connections constant.

My two facet model had good model fit and the results are consistent with other research in the area. While the gender facet in my model has predictive validity, it lacks construct validity. I did not develop these items to measure a clearly defined construct. However, the gender items appear to measure two dimensions of patriarchy: beliefs about traditional, largely patriarchal gender roles, which tend to be benign or benevolent; and also more hostile, sexist beliefs about the subordination of women in society.[157] Two items measure beliefs about the proper – but equally valued – roles that men and women should follow in the home and interpersonal relationships (HTRADFAM, EROUGH). Two measure more sexist beliefs about the subordinated position of women in politics and employment (HWMNRTS,

[157] *See generally* Peter Glick & Susan T. Fiske, *The Ambivalent Sexism Inventory: Differentiating Hostile and Benevolent Sexism*, 70 J. PERSONALITY & SOC. PSYCHOL. 491 (1996).

ESEXIST). One measures a more hostile belief about the value of femininity (HFEMININ).[158] That said, a fair label to place on these collected gender items and the construct that they measure is patriarchy.

To contrast with this model, I also ran Kahan's model, with the thirteen items loading on one latent variable (grid), using PID7 as a control variable, and using GUILTY as the outcome variable. The model fit was not good. The chi-square value for the overall model fit was significant (chi-square (88) = 800.65, $p < .001$), suggesting a lack of fit between the hypothesized model and the data. Examination of other indices showed an unacceptable model fit: RMSEA = .11, p(close) < .001; CFI = .84; TLI = .81; SRMR = .06.[1] The latent variable (grid) did predict the endogenous variable GUILTY (path coefficient = -.37, $p < .001$), as did PID7 (path coefficient = .14, $p = .001$).[159]

The structural equation modeling of the CCP's model is consistent with the exploratory factor analysis, in that the data does not support the CCP's one factor model. While Kahan's model had predicative validity (the grid latent variable was a significant predictor of agreement with the outcome judgment), the poor model fit suggests that the grid latent variable does not have construct validity. Further, the path coefficient for the gender facet in my two facet model (-.33) was about the same as the path coefficient for Kahan's grid factor (-.37), which also suggests that in this rape hypothetical, the gender items are doing the work. While "grid" has predictive validity, the predictive validity appears to be generated by one of the subfacets alone.

VII. DISCUSSION

The data suggests that patriarchy, not hierarchy, predicts outcome judgments in acquaintance rape cases. This is inconsistent with the Kahan's findings in *Culture, Cognition, and Consent*.

The data also suggests that the H-ES does not measure a single construct of "grid," but rather has subconstructs. This has important implications for other CCP studies. In particular, for the other CCP studies that involve racial issues and that used the H-ES,[160] a subfactor related to race may explain the variation in outcome judgments rather than a global construct of "grid." Structural equation modeling of those studies could reveal the answer.

---

[158] The race facet appears to measure race hierarchy, with three items that appear to measure agreement with programs like affirmative action and one that appears to measure hostile beliefs about the stereotyped black underclass.

[159] The correlation matrix and theoretical model are provided in the online appendix.

[160] *See Culture and Identity-Protective Cognition*, *supra* note 47 (includes gun risks that implicate race hierarchy and abortion risks that implicate gender hierarchy); *Self-Defensive Cognition*, *supra* note 55 (includes a vignette where a white man shoots black man in self-defense); *Whose Eyes*, *supra* note 80 (involves the use of police force on a fleeing suspect that implicates race hierarchy).

That said, a global construct of "grid" may exist. The high correlation between the two facets in my model suggests that there might be a higher-order construct, and an improved scale with reliable items might be able to measure that construct. One option for improving the measurement instrument would be to have all abstract or global questions,[161] like those used by John Jost and Erik Thompson when they measured a similar construct.[162] This would require the CCP to modify the "grid" definition somewhat to drop the facets.

Another option would be to measure the facets of this global "grid" construct. There could be four items each on gender, race, class (if that is a facet rather than a byproduct of hierarchical beliefs), lineage, and sexual orientation. The importance of each subfactor could be accounted for when calculating factor scores.[163] When the scale is used in projects that relate to a particular facet, like race or gender, researchers could then use structural equation modeling to see if the subfactors behave differently than the aggregated factor, as my model revealed is the case here, and then those could be reported separately from the aggregated score.

The CCP may also need to revise its construct definitions. When analysis suggests "that conceptualization of the target construct as, for example, a single bipolar dimension is countermanded by evidence that the two poles actually represent separate and distinct entities . . . revision of one's theoretical model may be in order".[164]

For the H-ES, the two factor loading revealed by the EFA and the poor model fit revealed by the CFA may be the result of the CCP's item development, where the CCP looked to items that were designed to measure features of Douglas' hierarchy and egalitarianism cultures.[165] Under her model, the constructs of hierarchy and egalitarianism represent cultures that form at the intersection of two dimensions (grid and group). As discussed above, while the CCP set "hierarchy" and "egalitarianism" on just one dimension (grid), the CCP definition of "egalitarian" includes a "group" component and many of the H-ES items appear to also tap into the "group"

---

[161] *See generally* DEVELLIS, *supra* note 13, at 74 ("There is general agreement in the social sciences that variables will relate most strongly to one another when they match with respect to level of specificity. Sometimes a scale is intended to relate to very specific behaviors or constructs, while at other times, a more general and global measure is sought.")

[162] John T. Jost & Erik P. Thompson, *Group-Based Dominance and Opposition to Equality as independent Predictors of Self-Esteem, Ethnocentrism, and Social Policy Attitudes Among African Americans and European Americans*, 36 J. EXPERIMENTAL SOC. PSYCHOL. 209 (2000). The items are listed in the online appendix.

[163] Haynes et al., *supra* note 20, at 244-45.

[164] Clark & Watson, *supra* note 13, at 312.

[165] Analysis of items used to measure the Douglas cultures reveals two dimensions. *See* Rippl, *supra* note 77, at 152.

construct. These items may actually be measuring features of the cultures that form at the intersection of the CCP's revised dimensions of "grid" and "group," which the CCP calls "hierarchical communitarian" and "egalitarian communitarian," rather than the dimensions themselves. Here, the CCP may need to replace "egalitarian" on the y-axis with "structureless," and then return "egalitarian" to a culture formed by the intersection of the two dimensions. The H-ES items would then need to be revised to clean out any aspects of the "group" dimension.

Last, based on the structure revealed by the EFA, the CCP's method of factor scoring is problematic. In *Culture, Cognition, and Consent*, Kahan reverse-scored the "E" items and appears to have calculated factor scores based on unweighted simple sums of the item scores (which he probably then standardized). Next, Kahan created poles along the grid construct by coding an observation as "hierarch" if the observation's score was in the top third of the simple sum distribution and as "egalitarian" if the observation's score was in the bottom third. Kahan then used those coded "hierarchs" and "egalitarians" as predictor variables.[166]

Assuming the factor structure revealed by this EFA is valid,[167] the EFA suggests that the H-ES measures at least two separate constructs and so all of the items should not be added together to arrive at a factor score. Instead, Kahan should have summed the scores for the "H" and "E" constructs and then used those scores as predictor variables.

This can lead to inaccurate results. The two are strongly, negatively correlated (r = -.66), but they are not perfectly correlated. Man respondents could score high (or low) on the "H" items and the "E" items alike. The "E" items are reverse-scored, which then moderates the resulting factor score.[168] This could put the respondent in the middle of the distribution and because Kahan only looked at the tails of the distribution, the respondent falls out of the study. This potentially masks the predictive value (or lack of predictive value) of the "H" and "E" constructs.

This calls into question the precise findings in *Culture, Cognition, and Consent*: "[T]he choice made regarding how factor scores are computed can significantly affect their quality as well as the outcomes of subsequent analyses in which the scores are used."[169] And, this may call into question

---

[166] *Culture, Cognition, and Consent*, *supra* note 1, at 776 fig.3, 777 fig.4, 780 fig.5, 785 fig.7, 786 fig.8, 792 fig.10.

[167] The "H" might measure the CCP's worldview of "hierarchical communitarian" and the "E" might measure "egalitarian communitarian." While EFA and CFA revealed that the H and E model did not fit well, the H and E model appears to fit better than the unidimensional model. With model tweaking or item refinement, this model might work.

[168] *See also* Olli, *supra* note 99, at 480.

[169] James W. Grice, *Computing and Evaluating Factor Scores*, 6 PSYCHOL. METHODS 430, 431 (2001) (citations omitted).

the precise findings of all the CCP studies that relied on this factor scoring method.[170]

While I have suggested that the H-ES has certain reliability and validity issues and that there are issues in the factor scoring, the general trends that the CCP spotted in these other studies probably still exist. In this particular study, the scale still had predictive validity and it likely has predictive validity in others. But, there is a loss in accuracy and we cannot be certain what the factor scores actually represent. It is difficult to understand the research domain and interpret the score measures in a meaningful way.[171]

VIII. CONCLUSION

The data suggests that beliefs related to patriarchy rather than worldviews related to hierarchy explain the guilt judgments in acquaintance rape cases. This also suggests that reform efforts do not need to be targeted at creating separate organizations that are independent of existing, hierarchical law enforcement, or military, or university, or religious organizations. Rather, targeted training on certain gender role beliefs (particularly beliefs about the ways that women should behave sexually) may create bias-free organizations. The findings of my study are consistent with other findings in the field, and those other findings show that those who hold traditional gender role or patriarchal beliefs subscribe to rape myths to a higher degree than those who do not hold those beliefs. When people use rape myths to resolve the discreet rape problem found in the legal file on their desk, they tend to find in favor of the man. Targeting those belief systems, rather than overhauling the organizations, may suffice.

---

[170] The CCP has used similar course factor scoring methods that combine the H and E items in other studies (although they may have occasionally used a least squares regression approach). *See Gender, Race, and Risk Perception*, *supra* note 47, at 12-13; *Self-Defensive Cognition*, *supra* note 55, at 35-36; *Whose Eyes*, *supra* note 80, at 860; *Outpatient Commitment Laws*, *supra* note 101, at 125; *Risks and Benefits of Nanotechnology*, *supra* note 101, at 89; *Affect, Values, and Nanotechnology Risk Perceptions*, *supra* note 51, at 11; *Who Fears the HPV Vaccine*, *supra* note 101, at 507; *Cultural Cognition of Scientific Consensus*, *supra* note 51, at 154; *Is Synthetic Biology Different?*, *supra* note 101, at 4, 10.

[171] Floyd & Widaman, *supra* note 130, at 287.