Journal of Business Ethics
https://doi.org/10.1007/s10551-018-3937-8

ORIGINAL PAPER

The Ethical Standards of Judgment Questionnaire: Development and Validation of Independent Measures of Formalism and Consequentialism

Ed Love[1] · **Tara Ceranic Salinas**[2] · **Jeff D. Rotman**[3]

Received: 22 November 2016 / Accepted: 1 June 2018

Abstract
The ethical frameworks of consequentialism and formalism predict moral awareness and behavior in individuals, but current measures either do not treat these frameworks as independent or lack sufficient theoretical underpinnings and statistical dependability. This paper presents the development and validation of a new scale to measure consequentialism and formalism that is well grounded in prior research. The Ethical Standards of Judgement Questionnaire (ESJQ) is validated via six studies (total $n > 2400$). Measurement items are developed in the first three studies, which also confirm the need to eliminate a unidimensional measure and evaluate these frameworks separately. The fourth study addresses discriminant validity and the two remaining studies provide insight into how consequentialism and formalism predict the degree to which behaviors are deemed acceptable by individuals in the context of consumer beliefs and religious beliefs. Suggested uses for the scale in both academia and organizations are presented.

Keywords Business ethics · Consequentialism · Deontology · Ethical decision making · Formalism · Religiosity · Scale development · Utilitarianism

Abbreviations

AVE	Average variance extracted
CFA	Confirmatory factor analysis
DUREL	Duke university religious index
EFA	Exploratory factor analysis
ESJQ	Ethical standards of judgement questionnaire
IR	Internal religiosity
MV-CES	Muncy–Vitell consumer ethics scale
MEP	Managerial ethical profile
MES	Multidimensional ethics scale
MEV	Measure of ethical viewpoints
MVP	Managerial values profile
NORA	Non-organized religious activity
ORA	Organized religious activity
SETA	Survey of ethical theoretic aptitudes

✉ Ed Love
Ed.Love@wwu.edu

✉ Tara Ceranic Salinas
tara@sandiego.edu

Jeff D. Rotman
Jeff.rotman@deakin.edu.au

[1] College of Business and Economics, Western Washington University, 516 High Street, Bellingham, WA 98225, USA

[2] School of Business, University of San Diego, 5998 Alcalá Park, San Diego, CA 92110, USA

[3] Deakin Business School, Deakin University, 221 Burwood Hwy, Burwood, VIC, Australia

Introduction

Ethical predispositions are foundational to how individuals reason. They shape both moral awareness (Reynolds 2006) and moral decisions (Brady and Wheeler 1996). Within organizations, they influence perceptions of justice (Schminke et al. 1997) and administrative policies (Alder et al. 2007). Reynolds and Ceranic (2007) point to ethical predispositions as, "[p]erhaps the most fundamental alternative for conceptualizing and measuring moral judgments" (p. 1611). Others have found their influence to transcend the boundaries of moral judgment (Love et al. 2015).

Descriptions of ethical predispositions are broadly based around two frameworks: consequentialism and formalism (Barak-Corren and Bazerman 2017). The distinctions between these approaches to decision making have been part of the moral landscape for thousands of years; some philosophers suggest that all ethics can be lumped into one of these two broad categories (Nozick 1981).

Published online: 20 June 2018

Individuals employing a consequentialist framework focus on the outcomes of their decisions and actions while those using a formalistic framework are motivated by the duty to follow rules and principles (Brady and Wheeler 1996; O'Shaughnessy 2002). Prior conceptualizations of consequentialism and formalism have utilized instruments that place these frameworks as opposite ends of a continuum (Brady 1985), but the more recent consensus is that they are distinct, yet related constructs (Brady and Wheeler 1996; Brady 1990; Burton et al. 2006; Conway and Gawronski 2013; Greene et al. 2008; Love et al. 2015; Pearsall and Ellis 2011; Reynolds and Ceranic 2007; Reynolds 2008). Conway and Gawronski (2013) note that,

> Although the traditional dilemma methodology has provided useful insights into moral psychology, it suffers from an important drawback. Participants must categorize a harmful action as either acceptable or unacceptable, thereby endorsing either the deontological or utilitarian principle. To behave in line with the deontological principle is to simultaneously behave in opposition to the utilitarian principle, and vice versa. Thus, the traditional approach confounds selecting one option with rejecting the other. This confound would be acceptable if the moral inclinations underlying overt deontological and utilitarian judgments were themselves inversely related (i.e., stronger inclinations of one kind are associated with weaker inclinations of the other kind). However, theorists have argued that deontological and utilitarian inclinations stem from conceptually distinct and functionally independent processes, thereby allowing for the possibility that both inclinations are active at the same time (Greene 2008). Indeed, moral dilemma research is predicated on the assumption that high-conflict dilemmas arouse conflict between the two inclinations (Greene et al. 2001; Koenigs et al. 2007) and that whichever inclination is stronger drives the behavioral response (i.e., judging harm as acceptable or unacceptable). Such conflict would not occur if the two competing inclinations were inversely related (p. 217).

Using a dissociation process to evaluate responses to moral dilemmas, Conway and Gawronski (2013) empirically demonstrate that formalism and consequentialism are independent constructs. They go on to conclude that prior research may have, "confounded the operation of the two inclinations by treating them as inversely related dimensions of a bipolar continuum" (p. 229).

Accordingly, use of a unidimensional measure for these distinct frameworks is problematic. As Love, Staton, and Rotman (2015) note,

> Deontology and utilitarianism are not opposing standards. If a deontologist is someone who adheres to principles when making decisions, an anti-deontologist is someone who ignores principles. This in no way suggests that they focus on outcomes instead of principles, as a unidimensional measure would suggest. It is as possible that the anti-deontologist would make decisions based on what requires the least cognitive effort as on what provides the most utility. Conversely, the anti-utilitarian need not adhere to principles when making decisions (p. 664).

The terms utilitarianism and consequentialism are often used interchangeably and the underlying philosophical concepts are closely related. In fact, the Stanford Encyclopedia of Philosophy refers to utilitarianism as the "paradigm case of consequentialism" (Sinnott-Armstrong 2015). However, the term consequentialism is better suited to describe a framework for decision making. The utilitarianism popularized by John Stuart Mill defines right action as that which maximizes well-being for the greatest number of people (Baron 2017; Rajczi 2016; Reynolds and Ceranic 2007). Mills' and Bentham's act utilitarianism is, philosophically speaking, the most common form of consequentialism, but while the emphasis of utilitarianism is on maximizing wellbeing, consequentialism refers to a belief that the rightness of a choice or act is a function of the consequences of that choice or act (Walsh 2015). Since the current research deals with standards of judgment, the term consequentialism is marginally preferred by the authors. Consequentialism, as used here, is equivalent to the functional definition of utilitarianism applied by Brady and Wheeler (1996). They define "utilitarian" as, "the tendency to assess ethical situations in terms of their consequences to people" (p. 928). We follow the common operationalization of consequentialism as referring to a belief that the rightness of a choice is a function of the results (consequences) of that choice, and that rightness is based upon the maximization of wellbeing.

Brady and Wheeler (1996) broadly define formalism, or deontology, as, "human tendency to assess ethical situations in terms of their consistent conformity to patterns or rules or some other formal features." (p. 928). In other words, an action is considered right or wrong based on the degree to which it conforms to moral rules for good behavior. For example, killing an innocent is bad regardless of the consequences (Conway and Gawronski 2013). Since they are defined by rules, correct action can be defined as "good" or "bad" in an absolute and normative sense.

Necessary Criteria for a New Measure

Based on the above, a new assessment of ethical standards of judgment should meet the following criteria:

1. Formalism and consequentialism are measured as independent constructs. Attempts to measure the two constructs on a single dimension risks confounding their operation (Conway and Gawronski 2013). 1a. Separate subscales for the two constructs are not sufficient. The subscales must show discriminant validity between the two constructs to minimize confounding effects.
2. The formalism subscale should be both reliable and valid, such that it captures the tendency of an individual to make ethical choices based on a set of rules which determine whether a decision or behavior is good or bad. Consequences are irrelevant.
3. The consequentialism subscale should be both reliable and valid and should capture the tendency of an individual to make judgments and evaluate ethical choices based on outcomes, or the net utility of the choice.
4. The measures should be applicable to a general population, at least within a Western cultural context.

To specifically address these criteria, this work develops, tests, and validates a new measure of consequentialist and formalist frameworks called The Ethical Standards of Judgment Questionnaire (ESJQ). Although ethical frameworks are difficult constructs to measure and quantify (Schminke 1997), we think it is both necessary and possible to improve on the current measures.

Existing Measures and Limitations

Many validated and generally accepted measures exist in the ethical decision-making literature; however, the two measures most relevant to the current work are the measure of ethical viewpoints (MEV) (Brady and Wheeler 1996) and the Survey of Ethical Theoretic Aptitudes (SETA) (Brady 1990). These measures have been utilized in prior business ethics research (Schminke 1997; Kumar et al. 1991; Kumar and Strandholm 2002; Middleton and Byus 2004; Smith et al. 2009; Love et al. 2015; Reynolds and Ceranic 2007) and address our desired focus on formalism and consequentialism, the two dominant ethical frameworks. However, each have limitations that the ESJQ aims to address while demonstrating greater reliability and validity. We also briefly review other measures from the ethical decision-making literature to present a comprehensive overview of the currently available options for researchers.

Measure of Ethical Viewpoints

The MEV is a two-part instrument that measures ethical reasoning with forced-choice vignettes and a 13-item character trait scale. The vignettes were written specifically to address a variety of topics and to yield responses that were not influenced by unintentional bias. In order to do this, Brady and Wheeler (1996) constructed four possible responses to each vignette, each representing a combination of consequentialist and formalist solutions and rationales. Their intent was to show that,

> A consequentialist should prefer both consequentialist actions and consequentialist reasons for the actions. The consequentialist should least prefer formalistic actions and formalistic reasons. Ethical theory would suggest that people make decisions that are consistent with their way of thinking, and that the rationale should tend to dictate the solution (Brady and Wheeler 1996, p. 929).

The second part of the instrument was a character traits measure, which provided a list of descriptors that one ethical viewpoint may appreciate or value over the other. As Brady and Wheeler note, the development of this trait scale was predicated on the assumption that "ethical predispositions might be associated with particular character traits and that persons would judge character traits to be important based on their ethical theoretic predispositions" (p. 931). Brady and Wheeler generated a list of twenty traits which they judged to be related to ethical predispositions. Each trait was rated on a 7-point Likert-type scale and these twenty items were evaluated by a sample of 141 financial services professionals with an exploratory factor analysis conducted on the results. In a multi-factor solution, the first factor explained 26% of the variance in responses to the twenty items, and the second factor explained 16.7% of the variance in the twenty items. Since no other factor explained more than 9% of the variance, they used the first two factors as the bases for their new trait scales.

Although Brady and Wheeler (1996) do not report specific factor loadings, they note that items they normally associate with utilitarian thinking—innovative, resourceful, effective, influential, results oriented, productive, and a winner—were "included" in the first factor, so they chose these items as the basis for a consequentialist subscale. Similarly, they note that the more formalistic items: principled, dependable, trustworthy, honest, noted for integrity and law-abiding were included in the second factor and so became the basis for a formalistic subscale.

Brady and Wheeler (1996) report acceptable reliability for both subscales, a finding that has been replicated many times in the literature. While they do not appear to conduct a confirmatory factor analysis in their scale development, subsequent users of the scale do. Schminke, Ambrose, and Noel (1997) report similar results to Brady and Wheeler's EFA and Reynolds and Ceranic (2007) find that, with the exception of the formalistic item "law-abiding," the items show acceptable convergent and discriminant validity. However, while Reynolds and Ceranic (2007) show a moderate

amount of correlation between the subscales ($r = 0.42$, $p < 0.01$), other published studies have reported much higher correlations between the measures. Schminke et al. (1997) report a subscale correlation of 0.53 ($p < 0.001$), Schminke (2001) separately reports the same correlation, and Alder, Schminke, Noel, and Kuenzi (2008) report a subscale correlation of 0.57 ($p < 0.01$). More recently, Letwin et al. (2016) report a high enough correlation between formalism and consequentialism ($r = 0.76$, $p < 0.01$) that a further investigation of discriminant validity is justified.

Separate from but related to the issue of discriminant validity, it is reasonable to question the construct validity of the MEV measures of formalism and consequentialism. In other words, do these subscales measure what they are intended to measure? Surprisingly, we find little evidence of validity checks in the literature beyond the initial study run by Brady and Wheeler (1996).

A careful examination of the items suggests that the scales may be interpreted as measuring constructs other than consequentialism and formalism. As Mudrack and Mason (2017) note, the consequentialist subscale measured by the MEV does not actually appear to be consistent with the typical conceptualization of utilitarianism in which the rightness of an action is determined by the degree to which it provides the greatest benefit to the greatest number of people (Mill 1863, as cited in; Mudrack and Mason 2017) and lends credence to this line of questioning.

In summary, despite the body of extant research using the MEV, it remains unclear whether this measure truly satisfies all of the criteria that we list above. Although the MEV utilizes distinct measures for formalism and consequentialism. The consistent correlation between the subscales found in previously published studies indicates that prior research cannot conclusively determine whether the MEV passes criterion 1a (discriminant validity). Since construct validity checks have not been included in prior research, the MEV's performance on criteria 2 & 3 (reliable and valid subscales for both formalism and consequentialism) must be considered indeterminate as well. The MEV does satisfy criterion 4 (generalizable) in that the scale is suitable for a general population of respondents.

Survey of Ethical Theoretic Aptitudes

The SETA Brady (1990) measures an individual's preference for a consequentialist or formalist framework. The instrument contains fifteen statements with two choices (A and B) for each. Example questions include, "Persons' actions should be described in terms of (A: good or bad; B: right or wrong)" and "It's more important to be (A: happy; B: worthy)." The SETA is based on measures of judgment standards and more accurately captures the categorical constructs, making it an attractive alternative to the MEV. Even so, the SETA has several limitations that need to be addressed.

First, the sentence completion task utilized by the SETA forces respondents to choose between non-opposite anchors. For example, the opposite of "Good or bad" is not "Right or wrong." This forced choice is suitable for capturing categorical preference for formalistic or consequential standards but does not capture the degree to which an individual favors either formalism or consequentialism. As stressed by Brady and Wheeler (1996), formalism and consequentialism are independent standards; a deontological outcome on the SETA could therefore be the result of high formalistic tendencies, low consequentialist tendencies, or practically any combination of the two—so long as the respondent's formalistic tendencies exceed their consequentialist tendencies. In other words, a deontological outcome on the SETA reveals little about a respondent's actual degree of formalism or consequentialism.

The SETA is limited in other ways as well. The items do not allow for any gradation, thus obliging subjects to select a phrase that may not necessarily align with their own answer to the question prompt and the lack of nuance between the choices leads to the loss of considerable information. Consider the item, "When making an ethical decision, one should pay attention to [one's conscience/others' needs, wants, and desires]." It is entirely reasonable that an individual would consider both one's own conscience and others' needs, wants, and desires when making an ethical decision. However, due to the design of the SETA, it is possible to assess the consistency of preference for one of the two standards, but not the degree of preference.

Additionally, there is lack of satisfactory reliability from a psychometric perspective. Schminke (1997) reported a Cronbach's alpha of 0.68 for that scale, and went on to suggest that the, "moderate reliability of the SETA may be responsible for the modest results" (p. 62). In multiple attempted replications, we were unable to match or surpass Schminke's reported alpha.

In a pretest ($n = 342$), a principal components analysis of the fifteen items of the SETA yielded only one component that explained more than 10% of the scale variance, and no component explained more than 16% variance. Six components had eigenvalues over one, none of which were readily interpretable. For example, the only two items with loadings of greater than 0.5 on the first component were "A nation should pay more attention to its - heritage, its roots: its future, its potential" and "It is of more value to societies to - follow stable traditions and maintain a distinctive identity: be responsive and adapt to new conditions as the world changes." No item loaded at more than 0.65 to any single component. Scale reliability in this pretest was poor ($\alpha = 0.564$).

Finally, the SETA has shortcomings in that scoring the fifteen items is a somewhat more cumbersome process than is used in more contemporary scales. One must "Count the odd-numbered questions to which you responded 'A,' Count the even-numbered questions to which you responded 'B,' add the two numbers together, and subtract 8 from the total" (Brady 1990, p. 211–213). This final score is then compared to a table based on which the respondent falls into one of the six categories ranging from "flaming consequentialist" to "ice-cold formalist." This scoring process has made the SETA less appealing than, say, the simpler MEV in research and, as previously noted, the resulting score does not provide any information about the *degree* of consequentialism or formalism preferred.

In summary, the SETA does not satisfy criterion 1 because it does not treat formalism and consequentialism as independent constructs. The SETA also does not satisfy criteria 2 & 3 in that it is not a consistently reliable measure, although the specific scale items do appear related to the underlying constructs. The cumbersome scoring process is a secondary concern that makes the SETA less attractive to use. Like the MEV, the SETA satisfies criterion 4 (generalizability).

These criticisms notwithstanding, the SETA's focus on formalistic and consequential judgements, as well as its brevity, make it appealing as the basis for a new scale. In the pretest, the SETA did correlate to a unidimensional validity check seven-point scale, "I tend to make decisions based on... My values, principles, and core beliefs, the outcomes that I expect my decisions will have on myself, my community, and the world in general." ($r = 0.312$, $p < 0.01$).[1]

Therefore, this work acknowledges the general utility of the SETA in capturing the constructs and employs it as the foundation for two more refined, shorter, reliable, and predictive scales.

Related Measures

Aside from the MEV and the SETA, there are a variety of scales that endeavor to measure ethical decision making. These scales largely attempt to evaluate ethical constructs other than formalism and consequentialism; however, Casali (2011) provides an excellent review, assessment, and explanation of the limitations of five of the most prominent of these measures: Defining Issues Test (DIT), Managerial Judgment Test (MJT), Ethics Position Questionnaire (EPQ), Managerial Value Profile (MVP), and Multidimensional Ethics Scale (MES).

Of these five measures, the MVP (Sashkin et al. 1997) and MES (Reidenbach and Robin 1988, 1990) categorize individuals according to their ethical decision-making principles. This approach is in line with the intent of the ESJQ, but each of these scales present limitations. The MVP's 12 forced-choice items provide a score from 0 to 8 for each of three ethical frameworks: utilitarianism, deontology, and social justice, and individuals are assigned an ethical 'profile' based on the scale with the highest score (Casali 2011). This approach does not account for the possibility of using a combination of ethical frameworks when making a decision, thus not satisfying criterion 1 and, similar to the SETA, a great deal of information is lost with forced-choice items. The MES is a multi-item scenario-based measure that asks respondents to determine the ethicality of presented situations based on ethical criteria derived from five ethical foundations: justice, relativism, utilitarianism, egoism, and deontology (Reidenbach and Robin 1988, 1990). More recent work by McMahon and Harvey (2007) points to limitations of the MES such as high cross-scale correlations and within rather than across scenarios analysis, which brings generalizability and applicability of the measure into question. Additionally, Reidenbach and Robin (1990) acknowledge that some components of their scale do not "correspond strictly to the normative philosophies and tend to disagree with several of the hypothesized relationships in recently developed models of ethical decision-making" (p. 649). Thus, MES cannot be said to satisfy criteria 2 and 3 (reliable and valid subscales for both formalism and consequentialism).

The remaining three measures, DIT, MJT, and EPQ, focus on "psychological aspects rather than moral theories" (Casali 2011, p. 487). The DIT and MJT have both been used extensively in ethics research; however, they both aim to categorize individuals according to their level of cognitive moral development (Kohlberg 1979) (thus not satisfying criteria 2 and 3) and utilize scenarios created to elicit an ethical response. This focus on moral development is different from the approaches taken by the MVP and MES; since these scales focus on categorizing individuals by their ethical principles (Casali 2011), again not satisfying criteria 2 and 3. The EPQ attempts to assess an individual's personal moral philosophy via measures of idealism and relativism (Forsyth, 1980), but like the SETA, this approach is limiting in that the measures are unidimensional and a false dichotomy is established between idealism and relativism (Casali 2011). In other words, the EPQ fails to satisfy criteria 1, 2, and 3.

[1] In a separate test ($n = 414$), the ESJQ measures of formalism and consequentialism presented later in this paper also correlated with this validity check ($r = -0.295$, $p < 0.01$ for formalism, $r = 0.133$, $p < 0.01$ for consequentialism).

Managerial Ethical Profile

To address the concerns with the five measures, Casali (2011) proposed the Managerial Ethical Profile (MEP). This measure, however, does not satisfy criterion 4; in that it focuses specifically on the managerial decision making of healthcare workers in Australia. The MEP includes multidimensional measures of four ethical foundations via two dimensions each (Economic egoism, Reputational egoism, Rule utilitarianism, Act utilitarianism, Virtue of self, Virtue of others, Act deontology, Rule deontology) which addresses the need for multidimensionality, but the items and underlying measures are highly contextual by design. For example, the measures of rule deontology include, "Ensuring that confidentiality is maintained at all times," "Maintaining a fair process at all times," and "Ensuring that the organization's 'duty of care' is maintained at all times." While these are likely contextually appropriate, few would argue that they provide a generalized measure of formalistic standards of judgment. Casali (2011) even highlights the limitations of the MEP and lack of replication; all which must be addressed in new measures.

In summary, there is a need for a new measure of formalism and consequentialism, one that combines the multidimensionality of the MEV with the conceptual validity of the SETA. Below we explain the process used to develop and validate the new scale using the SETA as its foundation. First, we restructured the items to eliminate non-bipolar anchors and to allow for levels of gradation by converting the statements into Likert-style items. Next, separate formalism and consequentialism scale items were identified using a factor analysis. Scales were tested using a confirmatory factor analysis and two-factor structures were compared: unidimensional and multidimensional. Finally, discriminant, convergent, and predictive validity were assessed.

Research Methodology

Study 1: Scale Item Development

The purpose of Study 1 is to identify items based on the SETA that measure an individual's tendency toward formalistic and consequentialist judgment since no interpretable factor structure was found in the SETA, even when the items were adjusted to allow for points of gradation between the anchors. This study makes two important changes to the SETA items. First, each of the items was divided into two five-point Likert items, one reflecting the formalism anchor and one reflecting the consequentialist anchor. For example, for the item, "Solutions to ethical problems are usually [some shades of gray/black and white]," respondents are asked how strongly they agree or disagree with the following statements: "Solutions to ethical problems are usually some shades of gray" and "Solutions to ethical problems are usually black and white." This resulted in 30 5-point Likert-type items, of which fifteen potentially measured formalistic tendencies and fifteen potentially measured consequentialist tendencies. Small changes were made to some items to make them more interpretable in the new format. For example, "When thinking through ethical problems, I prefer to [develop practical, workable alternatives/make reasonable distinctions and clarifications]" became "When thinking of ethical problems, I try to develop practical, workable alternatives."

Methodology

A sample of 301 individuals were recruited using Amazon's Mechanical Turk (MTurk) service. Each responded to the thirty-item online survey. The median age of respondents was between 26 and 34 years old. Median household income was between $40,000 and $50,000 per year. 48.5% of the respondents were single, never married, 87.4% of the respondents had at least some college education and 46.8% of the respondents were female.

Results

No interpretable single factor solution emerged in an exploratory factor analysis (EFA) using all 30 items, suggesting that no single underlying construct drives responses. This casts further doubt on the unidimensionality of ethical standards of judgment. However, analyzing the consequentialism and formalism items separately resulted in two clear single factor structures.

A single factor ($\lambda = 3.03$) explained over 20% of the variance of the consequentialist items, over twice as much variance as any other factor. Six items had a loading of 0.499 or higher on this factor and no other factor had more than two items load on it at this level, thus providing strong initial support for a single consequentialist construct (Table 1).

Based on this analysis, six statements ("When people disagree over ethical matters, I strive for workable compromises," "When thinking of ethical problems, I try to develop practical, workable alternatives," "It is of value to societies to be responsive and adapt to new conditions as the world changes," "Solutions to ethical problems usually are seen as some shade of gray," "When making an ethical decision, one should pay attention to others' needs, wants and desires," and "The purpose of the government should be to promote the best possible life for its citizens.") were selected for further analysis as the basis for a consequentialist scale. These six items show satisfactory scale reliability ($\alpha = 0.70$).

An analysis of the formalist items yielded similar results. Based on an EFA, a single factor explained over 21% of the item variance ($\lambda = 3.18$) and no other factor explained more than 14% of the variance and six items

Table 1 Exploratory factor analysis rotated component matrix consequentialism items

	Component				
	1	2	3	4	5
When people disagree over ethical matters, I strive for workable compromises	0.78				
When thinking of ethical problems, I try to develop practical, workable alternatives	0.675				
It is of value to societies to be responsive and adapt to new conditions as the world changes	0.658				
Solutions to ethical problems usually are seen as some shade of gray	0.54				
When making an ethical decision, one should pay attention to others' needs, wants and desires	0.537				
The purpose of the government should be to promote the best possible life for its citizens	0.499				
The aim of science should be to solve existing problems		0.762			
Unethical behavior is best described as causing some degree of harm		0.577			
Thinking of occupations, I could see myself as a benevolent legislator, seeking an improved life for all					
A person's actions should be described in terms of being good or bad			0.696		
Uttering a falsehood is wrong because depending on the results, it can lead to further problems			0.649		
A nation should pay the most attention to its future, its potential					
Whether a person is a liar or not is a matter of degree				0.844	
It is more important to be happy than worthy				0.658	
Being known as a person who has accomplished a lot and achieved much is important to me					0.803

Extraction method principal component analysis

Rotation method Varimax with Kaiser normalization

Items in Bold represent basis for consequentialist scale

Table 2 Exploratory factor analysis rotated component matrix formalism items

	Component			
	1	2	3	4
Solutions to ethical problems are usually black and white	0.700			
A person's actions should be described in terms of being right or wrong	0.699			
A nation should pay the most attention to its heritage, its roots	0.673			
Societies should follow stable traditions and maintain a distinctive identity	0.662			
Uttering a falsehood is wrong because it wouldn't be right for anyone to lie	0.625			
Unethical behavior is best described as a violation of some principle of the law	0.565			
Thinking of occupations, I could see myself as a wise judge, applying the law with fairness and impartiality		0.668		
Being known as a person who has integrity and is a person of principle is important to me		0.665		
When making an ethical decision, one should heed one's conscience		0.520		
It is more important to be worthy than happy				
The purpose of the government should be to secure justice and fair treatment			0.709	
The aim of science should be to discover truth			0.543	
When thinking through ethical problems, I try to make reasonable distinctions and clarifications				
Whether a person is a liar is a question of kind				0.854
When people disagree over ethical matters I strive for some points of agreement				

Extraction method principal component analysis

Rotation method varimax with Kaiser normalization

Items in Bold represent basis for formalism scale

loaded on this factor at 0.500 or higher. Three items loaded on a second factor at 0.500 or higher (Table 2).

The six statements that load on factor one ("Solutions to ethical problems are usually black and white," "A person's actions should be described in terms of being right or wrong," "A nation should pay the most attention to its heritage, its roots," "Societies should follow stable traditions and maintain a distinctive identity," "Uttering a falsehood is wrong

Table 3 The Ethical Standards of Judgement Questionnaire (ESJQ)

Item number	Item
C1	When people disagree over ethical matters, I strive for workable compromises
C2	When thinking of ethical problems, I try to develop practical, workable alternatives
C3	It is of value to societies to be responsive and adapt to new conditions as the world changes
C4	Solutions to ethical problems usually are seen as some shade of gray
C5	When making an ethical decision, one should pay attention to others' needs, wants and desires
C6	The purpose of the government should be to promote the best possible life for its citizens
F1	Solutions to ethical problems are usually black and white
F2	A person's actions should be described in terms of being right or wrong
F3	A nation should pay the most attention to its heritage, its roots
F4	Societies should follow stable traditions and maintain a distinctive identity
F5	Uttering a falsehood is wrong because it wouldn't be right for anyone to lie
F6	Unethical behavior is best described as a violation of some principle of the law

because it wouldn't be right for anyone to lie," and "Unethical behavior is best described as a violation of some principle of the law.") show acceptable scale reliability ($\alpha = 0.76$).

In summary, the SETA was deconstructed into 30 items, fifteen of which potentially represented consequentialist tendencies and fifteen of which represented formalist tendencies. Consistent with Brady's Janus-headed conceptualization of consequentialism and formalism as complementary traits (Brady 1985; Mudrack and Mason 2017), no satisfactory unidimensional scale was found based on the thirty items (or the fifteen original items on which they were based). EFAs were conducted on the consequentialism and formalism items separately, each resulting in a six-item proto-scale. These scales are further evaluated in Studies 2a, 3, and 4.

Study 2: Confirmatory Factory Analysis and Factor Structure Comparison

The purpose of Study 2 is to evaluate the fit of the two-scale model developed in Study 1 through a confirmatory factor analysis (CFA). Although not supported in Study 1, a single factor model was also tested in order to evaluate the relative fit of the two-factor model. The items tested in the models are summarized in Table 3.

Method

A sample of 300 responded to the twelve items as part of a broader survey of ethical decision making.[2] As with Study 1, respondents were recruited through Amazon's MTurk. The median age of the respondents was between 35 and 44 years. Sixty-one percent had at least some college experience, 58% were female, 50% were single, 31% were married, 12% were living with a partner, and 9% were either divorced, widowed, or separated.

Results

A confirmatory factor analysis of the two-factor structure using the lavaan package in R indicated a reasonable model fit (cmin/df = 2.6, RMSEA = 0.07, CFI = 0.91, SRMR = 0.07). All parameters loaded significantly on their expected factors ($p < 0.01$ in all cases, see Table 4). More importantly, this two-factor solution provided a clearly superior fit to the single factor model (cmin/df = 6.7, RMSEA = 0.138, CFI = 0.685, SRMR = 0.137).

Table 4 Study 2 parameter estimates, 2 factor CFA

	Estimate	SE	C.R.
Consequentialist			
C6	1		
C5	0.886**	0.183	4.848
C4	0.593**	0.15	3.964
C3	0.751**	0.161	4.668
C2	1.266**	0.205	6.181
C1	1.225**	0.227	5.405
Formalist			
F6	1		
F5	0.963**	0.164	5.882
F4	0.96**	0.165	5.819
F3	0.911**	0.166	5.478
F2	1.046**	0.174	6.023
F1	0.927**	0.158	5.86

**$p < 0.01$ level (2-tailed)

[2] The survey also included psychometric measures and moral dilemmas, but these questions appeared after the items discussed in this paper and are not relevant to the current research.

Table 5 Study 3 parameter estimates, 2 factor CFA

	Estimate	SE	C.R.
Consequentialist			
C6	1		
C5	0.599**	0.202	2.97
C4	0.421†	0.237	1.78
C3	0.774**	0.207	3.75
C2	0.826**	0.257	3.22
C1	1.022**	0.319	3.21
Formalist			
F6	1		
F5	0.89**	0.2	4.45
F4	1.009**	0.232	4.36
F3	0.722**	0.25	2.89
F2	0.757**	0.25	3.03
F1	0.339*	0.137	2.48

**$p < 0.01$ level, *$p < 0.05$, †$p < 0.10$ (2-tailed)

These results provide strong support for the two-factor model developed in Study 1. Expanding on the results of the previous study, the CFAs suggest that consequentialism and formalism are better measured as two separate constructs than as a single construct. Also, the six consequentialism items and the six formalism items consistently align with the underlying constructs.

Study 3: Replication of Confirmatory Factory Analysis Using In-Person Sample

Study 3 replicates the findings of study 2 using an in-person student sample rather than an online MTurk sample. This provides greater external validity for our findings and evidence that our findings are not due to a possible response bias from an online pool.

Method

Two hundred and thirty-seven students from a large Canadian university participated in this study. Seven individuals failed to complete the survey, leaving a total of 230 (52.3% female). Participants completed the same ESJQ items as in study 2.

Results

Results of a confirmatory factor analysis of student responses, conducted using AMOS, provide good support for the two-factor model developed in Study 1 and tested in Study 2. The two-factor model appears to fit the data well (cmin/df = 1.8, RMSEA = 0.06, CFI = 0.90, GFI = 0.95). As shown in Table 5, all factors loaded on their expected factors ($p < 0.08$ for "Solutions to ethical problems usually are seen as some shade of gray," $p < 0.05$ for all other items).

Study 4

As noted previously, an analysis of prior research raises questions about whether the MEV satisfies each of the criteria that we propose for a measure of ethical orientations. Specifically, the high reported correlations between the measures of formalism and consequentialism leads to concerns about discriminant validity between the measures (criterion 1a). Also, questions about the specific subtrait items and their relationships to consequentialism lead to concerns about construct validity (criterion 3). In study 4, we assess the discriminant validity of both the ESJQ and the MEV. Doing so in a single study provides the opportunity to compare results between the measures. As noted above, a consequentialist should consider the consequences of a solution when evaluating ethical judgements, yet prior research has raised concerns about the degree to which the MEV trait measure actually captures this tendency. We therefore developed a validity check for the consequentialist constructs based on these conceptualizations.

Method

One hundred and fifty-two MTurk respondents completed the ESJQ and the MEV. Respondents also completed a consequentialist validity check ("Solutions to ethical problems should depend on the consequences") in order to establish construct validity. Responses to the validity checks were measured on five-point likert scales ("Strongly Disagree" to "Strongly Agree").

Fifteen respondents failed an attention check[3] and were excluded from the analysis. The median age of included respondents was 33. Forty-five percent of the sample was female, 29% were married, 44.5% single (never married), 17% divorced, widowed or separated; 16% living with partner, and 34% had children.

Results

Reliability for both versions of the consequentialism and formalism scales was adequate ($\alpha = 0.783$ and 0.750, respectively, for the ESJQ scales, $\alpha = 0.848$ and 0.904, respectively, for the MEV scales).

[3] "For this study, and for most of the research we conduct, it is very important for us to know whether you tend to read the instructions you are given. If you are reading these instructions, please just skip the following question and proceed to the next page. What type of products experiences were you asked about in this study?" Respondents who did not skip the question were excluded.

Table 6 Correlations between ESJQ measures and ESJQ scale items

Item	Mean	S.D.	1. ESJQ consequentialism	2. ESJQ formalism
1. ESJQ consequentialism	3.810	0.616	1	
2. ESJQ formalism	3.051	0.699	−0.068	1
3. When people disagree over ethical matters, I strive for workable compromises	3.693	0.854	0.721**	−0.004
4. When thinking of ethical problems, I try to develop practical, workable alternatives	3.774	0.813	0.689**	0.165
5. It is of value to societies to be responsive and adapt to new conditions as the world changes	3.905	0.839	0.653**	−0.169*
6. Solutions to ethical problems usually are seen as some shade of gray	3.613	0.980	0.668**	−0.222**
7. When making an ethical decision, one should pay attention to others' needs, wants and desires	3.847	0.839	0.726**	−0.049
8. The purpose of the government should be to promote the best possible life for its citizens	4.029	0.992	0.713**	0.019
9. Solutions to ethical problems are usually black and white	2.387	1.100	−0.378**	0.593**
10. A person's actions should be described in terms of being right or wrong	3.168	0.967	-0.070	0.745**
11. A nation should pay the most attention to its heritage, its roots	2.971	1.029	0.049	0.706**
12. Societies should follow stable traditions and maintain a distinctive identity	3.197	1.042	0.001	0.711**
13. Uttering a falsehood is wrong because it wouldn't be right for anyone to lie	3.445	1.028	0.055	0.570**
14. Unethical behavior is best described as a violation of some principle of the law	3.139	1.119	0.079	0.685**

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

Addressing the issue of discriminant validity, we note that while the ESJQ subscales are not significantly correlated ($r = -0.068$, n.s.), the significant correlation observed between the MEV subscales ($r = .631$, $p < 0.01$) falls within the range reported in previously published research.

Since the ESJQ subscales are essentially uncorrelated, discriminant validity between the subscales is not a practical concern. However, we conducted the same two checks of discriminant validity on both sets of measures for the sake of consistency. First, we used the Fornell-Larker test of discriminant validity (Fornell and Larcker 1981), in which the average variance extracted (AVE) by a factor analysis of the scale items is compared to the square of the correlation between the scale and some other measure (another scale, for example). The AVE represents "the average amount of variance that a construct explains in its indicator variables relative to the overall variance of its indicators" (Henseler et al. 2015, p. 116–117). Where the AVE from a scale's items is greater than the square of the correlation coefficient (r^2) between the scale itself and the measure with which it is being compared, the scale is considered to be discriminant from the other measure.

Using SPSS, we calculated the AVE figures based on the factor loadings in the rotated component matrix (varimax method). For ESJQ consequentialism, the factor analysis of the items produced only one factor with an eigenvalue of greater than one. AVE for the scale items was 0.485, which exceeds the r^2 of 0.005 between ESJQ consequentialism and ESJQ formalism. Similarly, one factor emerged from the ESJQ formalism item and the AVE for the ESJQ formalism items was 0.453, also exceeding the r^2 of 0.005.

MEV consequentialism and formalism correlated at 0.631, resulting in an r^2 of 0.398. A factor analysis of the consequentialism items resulted in a two-factor solution and an AVE of 0.396. Since this AVE falls below the r^2 threshold, we find that MEV consequentialism lacks discriminant validity. Although the MEV formalism scale fared better, with a single factor solution and an AVE of 0.691, the MEV cannot be said to satisfy criterion 1a.

Recent criticism of the Fornell-Larker (Henseler et al. 2015) method has observed that variance-based Structural Equation Modeling methods tend to over-estimate the AVE. Over-estimation of AVE could lead to scales being considered discriminant when, in reality, they lack discriminant validity. However, since the MEV consequentialism subscale was found to lack discriminant validity and the ESJQ subscale AVEs were nowhere near the discriminant validity thresholds, we deem this test to be sufficient.

We also tested discriminant validity by examining the cross loadings of the scale items via an assessment of item-level discriminant validity (Gefen and Straub 2005; Henseler et al. 2015). Rather than offering any specific threshold, discriminant validity is considered lacking where items from one scale correlate highly with a measure of a different construct. Essentially, we should expect items from the consequentialism subscale to have a low correlation with formalism, and items from the formalism subscale to have a low correlation with consequentialism.

We provide the correlation coefficients between ESJQ consequentialism, ESJQ formalism, and the ESJQ consequentialism and formalism items in Table 6. The consequentialist subscale consists of items 3–8, the formalist subscale

Table 7 Correlations between the MEV measures and the MEV scale items

Item	Mean	SD	1. MEV Consequentialism	2. MEV Formalism
1. MEV consequentialism	5.197	1.114	1	
2. MEV formalism	5.850	1.193	0.631**	1
3. Innovative	5.125	1.612	0.717**	0.368**
4. Resourceful	5.796	1.318	0.691**	0.571**
5. Effective	5.810	1.179	0.726**	0.649**
6. Influential	4.221	1.837	0.787**	0.430**
7. Results oriented	5.285	1.455	0.766**	0.514**
8. Productive	5.662	1.346	0.754**	0.705**
9. A winner	4.460	1.906	0.688**	0.173*
10. Principled	5.693	1.433	0.488**	0.772**
11. Dependable	6.015	1.366	0.570**	0.877**
12. Trustworthy	6.182	1.341	0.560**	0.884**
13. Honest	6.185	1.350	0.499**	0.862**
14. Noted For integrity	5.723	1.469	0.509**	0.839**
15. Law abiding	5.321	1.658	0.509**	0.754**

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

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

Table 8 Correlations between ESJQ measures, MEV measures, and validity checks

Item	Mean	S.D.	1.	2.	3.	4.	5.
1. ESJQ consequentialism	3.810	0.616	1				
2. ESJQ formalism	3.051	0.699	−0.068	1			
3. MEV consequentialism	5.197	1.114	0.220**	0.403**	1		
4. MEV formalism	5.850	1.193	0.372**	0.349**	0.631**	1	
5. Consequentialism check	3.504	0.925	0.322**	0.119	0.098	−0.017	1

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

consists of items 9–14. All items correlate to their intended scale at $r = 0.570$ or higher and that no item correlates to the unintended scale at a magnitude of greater than $r = 0.378$ thus confirming the discriminant validity of the two scales.

The pattern of correlations is somewhat different for the MEV measures and the MEV items (here, the consequentialism subscale consists of items 3–9 and the formalism subscale consists of items 10–15). Every item in correlates significantly to both subscales. We observe unintended correlations as high as 0.705, and ten of the thirteen items correlate with the unintended scale at $r = 0.488$ or higher (Table 7).

Our assessment of construct validity yielded satisfactory results for ESJQ consequentialism, with the consequentialism subscale predicting the consequentialism check ($r = 0.322$, $p < 0.01$). The MEV consequentialism subscale, however, did not predict the consequentialism check ($r = 0.098$, n.s.). Based on these results, the ESJQ appears to satisfy criteria 2 and 3, while the MEV satisfies criterion 2 but does not satisfy criterion 3 (reliable and valid consequentialism subscale) (Table 8).

Discussion

In study 4, we evaluated the discriminant validity of the ESJQ and the MEV scales. We found that ESJQ shows satisfactory discriminant validity but that the MEV does not. Similarly, while the ESJQ consequentialism scale shows satisfactory construct validity, the MEV does not. The traits measured in the MEV formalism subscale do appear to measure the intended ethical orientation, but the same cannot be said for the traits measured in the MEV consequentialism subscale. This scale does not predict the degree to which a respondent believes that ethical problems should depend on consequences, a key aspect of the consequentialist orientation. Further, the measure cannot be said to discriminate between consequentialism and formalism. As previously noted, this lack of discriminant validity creates a troublesome confound in empirical findings.

Table 9 Correlations between ESJQ measures and Muncy–Vitell consumer ethics survey items

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. Consequentialism	1								
2. Formalism	−0.062	1							
3. Actively benefitting from illegal activities	−0.203**	−0.147**	1						
4. Passively benefitting	−0.059	−0.269**	0.734**	1					
5. Benefiting from deceptive or questionable but legal activities	−0.017	−0.229**	0.719**	0.783**	1				
6. No Harm/No Foul	0.172**	−0.278**	0.335**	0.543**	0.510**	1			
7. Downloading/Buying counterfeit goods	0.035	−0.182**	0.577**	0.654**	0.626**	0.650**	1		
8. Recycling	0.360**	−0.160**	−0.274**	−0.080	−0.042	0.325**	−0.026	1	
9. Doing good	0.280**	−0.230**	−0.297**	−0.076	−0.065	0.335**	−0.021	0.678**	1

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

Study 5

To further evaluate the predictive validity and usefulness of the ESJQ measures, we assessed the impact of formalism and consequentialism on ethical perceptions with the Muncy-Vitell Consumer Ethics Scale (MV-CES) (Vitell and Muncy 2005). The MV-CES uses 31 items to measure seven aspects of consumer ethical beliefs. These aspects are "Actively benefitting from illegal activities," "Passively benefitting," "Benefiting from deceptive or questionable but legal activities," "No Harm/No Foul," "Downloading/Buying counterfeit goods," "Recycling," and "Doing good."

Since each item is measured on a five-point scale in which 1 represents "Strongly believe it is wrong" and 5 represents "Strongly believe it is not wrong," higher scores represent greater ethical acceptability of the types of decisions measured in each aspect of ethical belief. For example, a higher score in the dimension "Actively benefitting from illegal activities" indicates a greater tendency to believe that it is acceptable to benefit from such activities.

We expect consequentialism to be positively associated with aspects of ethical judgement that are linked to the outcomes or consequences of such aspects. These include "No Harm/No Foul" (since such activities do not cause harm, they should be viewed more favorably to consequentialists), "Recycling," and "Doing Good." Formalism, on the other hand, should be associated with aspects representing violations of rules or principles. These include "Actively benefitting from illegal activities," "Passively benefitting," and "Benefiting from deceptive or questionable but legal activities."

Method

Five hundred and one MTurk respondents completed the ESJQ and the MV-CES. The median age of respondents was 32. Forty-eight percent of the sample was female, 43% were married, 37% single (never married), 9% divorced, widowed or separated; 11% living with partner and 42% had children.

Results

A confirmatory factor analysis conducted in AMOS on the ESJQ items indicated acceptable model fit (GFI = 0.925). Reliability for both the consequentialism and formalism scales were adequate (α = 0.76 and 0.72, respectively). Consequentialism was significantly correlated with the outcome-related aspects of the MV-CES, whereas formalism was significantly correlated with the rule-based aspects of the scale (see Table 9).

Discussion

Other, not specifically predicted relationships revealed in this study may provide greater insight into the systems of judgement employed by consequentialists and formalists. For example, we observe that formalism negatively correlates with the perceived acceptability of benefitting from illegal activities (suggesting that formalists view benefitting from such activities as unacceptable). Also, while consequentialism is negatively correlated with actively benefitting from illegal activities (in other words, consequentialists tend to view this as unacceptable), consequentialism is not correlated with passively benefitting from illegal activities. We might infer that agency appears to matter to consequentialists. Similarly, the negative correlation between formalism and the acceptability of "Downloading/Buying counterfeit goods" provides insight into how these actions are evaluated. While one may view buying counterfeit goods as an act that harms the legitimate purveyors of such goods, these results suggest that a negative response to such purchases is more related to their illegality than their harm (Rotman et al. 2018).

The importance of treating formalism and consequentialism as orthogonal is further illustrated in the relationship between these orientations and "Actively benefitting from illegal activities." Since both orientations negatively correlate with this aspect of consumer ethics, it is unlikely that a unidimensional scale (such as the SETA) would provide any information about these relationships.

The acceptability of actively benefitting from illegal activities is of particular interest to managerial ethicists. Although the items used to measure this dimension are framed in the context of consumer judgment, they are likely to reflect employee judgment as well. Using the Hayes PROCESS tool in SPSS (Hayes 2012), we conducted a moderation analysis in order to better understand the interaction between formalism, consequentialism, and the acceptability of actively benefitting from illegal activities. As in the correlation analysis, the moderation analysis ($R^2 = 0.075$) showed that both formalism ($\beta = -0.216$, $p < 0.01$) and consequentialism ($\beta = -0.288$, $p < .01$) predicted the consumer ethics measure. The results also showed a positive interaction between the two terms ($\beta = 0.153$, $p < 0.05$), suggesting that the influence of the two standards is not additive.

Study 6: ESJQ and Religious Affiliation

The purpose of Study 6 is to assess the predictive validity of the two scales by exploring the relationship between religious affiliation and the ethical orientations. Barak-Corren and Bazerman (2017) highlight the well-established relationship between formalism and religiosity, noting that, "Research in moral psychology has found that religious decision-makers generally tend to form rule-based (deontological) judgments rather than outcome-based (utilitarian or consequentialist) judgments" (Barak-Corren and Bazerman 2017).

This statement is slightly problematic in that it argues for a categorical preference for formalistic judgments over consequentialist judgments among religious decision-makers. As discussed above, however, categorical preference provides little information regarding the importance an individual places on either category. This statement could imply either a highly formalistic approach or a highly anti-consequentialist approach to decision making (or both). Barak-Corren and Bazerman (2017) also offer several illustrative examples of religious individuals and organizations making decisions that are not only formalist but also arguably anti-consequentialist.

Method

One-thousand and nine respondents were recruited via Amazon MTurk to complete a survey on ethics, charitable giving, and religion. 119 failed an attention check,[4] leaving 890 respondents of which the median age was 36. Fifty-seven percent of respondents were female, 41% reported having children, 38% of the sample identified as single and 35% were married (the remaining respondents were either divorced, widowed, separated, living with partner, or preferred not to answer).

After responding to the ESJQ items, participants responded to a series of ethical problems unrelated to the current work. Respondents were asked "which of these best describes your religious affiliation" from a randomized list that included Christian, Jewish, Muslim, Buddhist, Hindu, Other faith, Unaffiliated faith, Atheist or Agnostic, and Don't Know/Prefer not to answer. Next, they were asked to complete the 5-item Duke University Religion Index (DUREL), which measures internal religiosity (IR), organized religious activity (ORA), and non-organized religious activity (NORA) (Koenig and Büssing 2010). The overall mean consequentialism and formalism scores were 3.89 and 3.10, respectively.

Results

A confirmatory factor analysis of the ESJQ items conducted in AMOS indicated acceptable model fit (GFI = 0.938). Reliability for both the consequentialism and formalism scales were adequate ($\alpha = 0.70$ and 0.76, respectively). Of the 890 respondents, religious affiliation was as follows: 429 Christian (48%), 16 Jewish (2%), 8 Muslim (1%), 15 Buddhist (2%), 8 Hindu (1%), 19 other faith (2%), 74 unaffiliated faith (8%), 284 Atheist or Agnostic (32%), 37 don't know/prefer not to answer (4%).

Respondents identifying as members of three of the major religious traditions: Christian, Muslim, and Hindu all employed formalistic judgment in comparison to those identifying as Jewish, Atheist/Agnostic or members of other faiths who preferred a non-formalistic approach (see Table 10; Fig. 1).

Additionally, formalism was positively correlated with IR, ORA, and NORA while consequentialism was negatively correlated with the same measures (See Table 11). In order to better understand the relationship between these constructs, we regressed mean-centered formalism, consequentialism, and an interaction variable (formalism x consequentialism) on each of the three measures of on IR, OR, and NORA. The results, shown in Table 12, indicate

[4] Respondents were given the following attention check: "For this study, and for most of the research we conduct, it is very important for us to know whether you tend to read the instructions you are given. If you are reading these instructions, please just skip the following question and proceed to the next page." Those that did not skip the subsequent question failed the check.

Table 10 Mean-centered ESJQ measures by religious affiliation

Which of these best describes your religious affiliation	Consequentialism			Formalism			Mean (zero-centered) religiosity		
	Mean	95% CI		Mean	95% CI		Internal religiosity	Organized religious activity	Non-organized religious activity
	(Zero-centered)	Lower Bound	Upper bound	(Zero-centered)	Lower bound	Upper bound			
Christian ($n=429$)	−0.10	−0.15	−0.05	0.29	0.22	0.35	0.94	0.90	0.89
Jewish ($n=16$)	−0.05	−0.30	0.20	−0.28	−0.59	0.05	0.43	−0.03	0.63
Muslim ($n=8$)	0.11	−0.28	0.51	0.50	0.25	0.78	1.52	1.29	2.25
Buddhist ($n=15$)	0.07	−0.15	0.26	−0.05	−0.32	0.22	0.29	0.23	0.88
Hindu ($n=8$)	−0.12	−0.50	0.23	0.27	−0.12	0.64	0.52	1.29	0.00
Other faith ($n=19$)	0.25	0.02	0.49	−0.44	−0.76	−0.14	0.69	−0.08	0.96
Unaffiliated faith ($n=74$)	0.04	−0.08	0.17	−0.10	−0.26	0.05	−0.06	−0.79	0.20
Atheist or Agnostic ($n=284$)	0.16	0.10	0.21	−0.39	−0.47	−0.32	−1.51	−1.16	−1.49
Don't know/Prefer not to answer ($n=37$)	−0.25	−0.44	−0.06	0.03	−0.18	0.21	−0.34	−0.53	−0.76
Total ($n=890$)	0.00	−0.03	0.03	0.00	−0.05	0.04	0.00	0.00	0.00

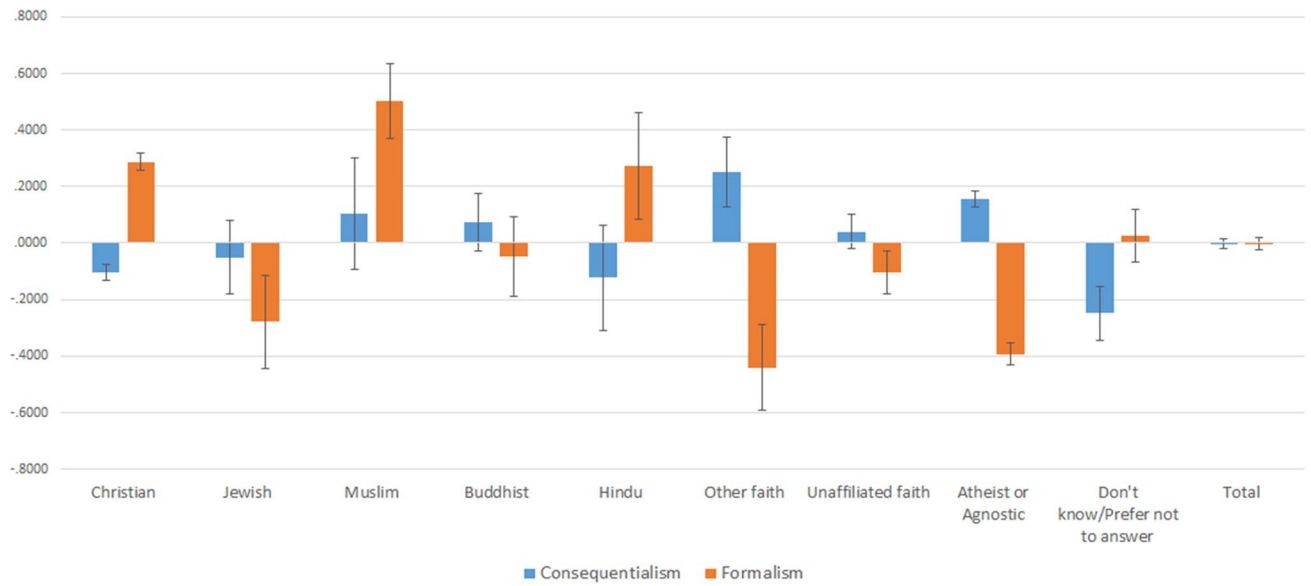

Fig. 1 Mean-centered ESJQ measures by religious affiliation

Table 11 Correlations between ESJQ measures, religiosity, and religious activity

	Mean	SD	1.	2.	3.	4.	5.
1. Consequentialism	3.89	0.54	1				
2. Formalism	3.10	0.70	−0.182**	1			
3. Internal religiosity	2.72	1.44	−0.157**	0.452**	1		
4. Organized religious activity	2.34	1.54	−0.181**	0.382**	0.663**	1	
5. Non-organized religious activity	2.6	1.84	−0.105**	0.352**	0.765**	0.567**	1

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

Table 12 Analysis of formalism as mediator of consequentialism-religiosity

	Regression Dependent Measure		
	Internal religiosity	Organized religious activity	Non-organized religious activity
R^2	0.21	0.159	0.131
ESJQ measure			
Formalism	0.92**	0.79**	0.96**
Consequentialism	-0.20*	−0.34**	−0.10 (n.s.)
Formalism by consequentialism	−0.03 (n.s.)	−0.07 (n.s.)	−0.33*

*p < 0.05, **p < 0.01

that formalism positively predicts each type of religiosity even controlling for consequentialism. Consequentialism no longer predicts NORA in these models.

Discussion

This study provides initial evidence of a more nuanced relationship between ethical standards, faith, and religiosity. Interestingly, our findings may contradict the notion that individuals who self-identify as belonging to a religious affiliation are more rule-based (formalist) (Barak-Corren and Bazerman 2017), as the participants in our sample who identified as Jewish, other or unaffiliated faiths preferred non-formalistic judgment.

Prior research cited by Barak-Corren and Bazerman (2017) similarly suggests both increased formalism and reduced consequentialism among religious individuals (Banerjee et al. 2010; Piazza 2012; Piazza and Sousa 2013). Piazza and Landy (2013) argue convincingly that the reason for increased deontological belief among the religious may be attributed to their meta-ethical belief that "morality is founded on God's moral authority." However, while a substantial amount of recent research has explored the relationship between religion and morality, none of this empirical work treats consequentialism and formalism as distinct, independent constructs.

It is beyond the scope of the current research to propose any specific causal relationship between formalism and religiosity, or between consequentialism and religiosity. Our purpose is not to determine whether increased/reduced formalism or consequentialism leads to greater/less religiosity, or vice versa, or if some third factor explains these relationships, but rather to show how formalism and consequentialism, as measured using the ESJQ, correlate to religiosity. Secondarily, we propose that affiliation with specific faiths is linked to systematic differences in ethical standards, and that these differences are captured in the ESJQ.

Discussion, Limitations, and Implications

These six studies offer a variety of evidence to support the contention that consequentialism and formalism are, indeed, distinct constructs. The first study establishes that the two ethical orientations are best captured as independent constructs rather than as a single underlying factor (as in the SETA) (Study 1), and this two-factor model was confirmed via CFA (Studies 2 & 3). In that, the MEV also treats the constructs as independent; this two-factor model is not new; however, our measures offer a substantial improvement over the MEV, as the items are based on ethical standards of judgment. Study 4 tests the discriminant validity of both the ESJQ and the MEV and shows that the MEV fails a validity check for the consequentialist subscale. Study 5 utilizes the MV-CES and the ESJQ to better understand how consequentialism and formalism impact an individual's ethical beliefs as they pertain to benefitting from various ethical and unethical behaviors. Finally, in Study 6, we observe that one's religious affiliation is strongly related to these standards of judgment in that religiosity is positively linked to formalism and negatively linked to consequentialism. Formalism was strongly linked to an affiliation with the Christian, Muslim, and Hindu faiths, and negatively linked to the Jewish faith, "other" faiths, and to atheism/agnosticism. We note that Judaism is a cultural affiliation as well as a religious one, which may explain the difference between Jews and respondents of other declared faiths.

The generalizability of this work may be limited by the use of online samples. However, research on MTurk shows

that these samples are comparable to respondents on other platforms (Huff and Tingley 2015) and the student sample used in study 3 helps to alleviate some of these concerns. We also recognize that there are a number of other measures of ethical decision-making constructs that exist, but examining further measures was beyond the scope of this paper.

From a scholarly perspective, the reliability, validity, and ease with which the ESJQ can be administered lends it to being applied in a variety of settings. Since a sufficiently reliable measure does not currently exist to test individuals' ethical judgements, we hope that the ESJQ will prove exceedingly useful to business ethics researchers.

Similarly, the ESJQ can help to inform a growing body of research in moral psychology. Specifically, recent research in moral psychology has been exploring the underlying mechanisms of individual judgments and responses to moral dilemmas (e.g., 'trolley problems'; Paxton, Unger & Greene, 2012; Conway and Gawronski 2013; Hütter et al. 2016; Duke and Bègue 2015). To examine whether the ESJQ can help inform these responses, a post-test was conducted. In addition, we examined whether the ESJQ provided additional predictive validity beyond the SETA and the MEV. Two hundred participants of which 16 failed the attention check, ($M_{age} = 36.3$, 59% female) completed the ESJQ, the SETA, the MEV, as well as five moral dilemmas (Paxton et al. 2012; Green et al. 2008) in which participants read about a moral dilemma and then were asked how ethical the action was (1 = not all ethical, 7 = very ethical). The dilemmas were coded so that higher values reflected the consequentialist response and a mean of the five measures was computed ($\alpha = 0.89$).

A regression demonstrated that both the ESJQ-Consequentialism ($\beta = 0.17$, $p < 0.05$) and ESJQ-Formalism variable ($\beta = -0.29$, $p < 0.001$) uniquely predicted the responses on the moral dilemmas. A similar regression revealed that although the MEV-Formalism variable predicted the moral dilemmas ($\beta = -0.24$, $p < 0.01$), the MEV-Consequentialism variable did not ($\beta = 0.12$, *ns*). Importantly, while the SETA predicted responses on the moral dilemmas ($r = 0.27$, $p < 0.001$), when the SETA, ESJQ-consequentialism, and ESJQ-formalism were all regressed on the moral dilemmas, consequentialism ($\beta = 0.17$, $p < 0.05$) and formalism ($\beta = -0.30$, $p < 0.001$) remained unique predictors, while the SETA became non-significant ($\beta = 0.01$, *ns*). The same pattern of results is seen when the MEV is also included in the regression, with the ESJQ variables remaining predictive ($\beta_{cons} = 0.17$, $p < 0.05$ and $\beta_{form} = -0.31$, $p < 0.001$), while the MEV variables were not ($\beta_{cons} = 0.13$, *ns* and $\beta_{form} = -0.14$, *ns*). Lastly, examining the relative R^2, the ESJQ, utilizing both consequentialism and formalism explain a significantly greater amount of variance ($R^2 = 0.15$), compared to the SETA ($R^2 = 0.07$) or the MEV ($R^2 = 0.05$). Thus, examining the relative moral predilections of the ESJQ provides both additional predictive ability and a better theoretical understanding of the underlying processes of moral judgment beyond the SETA and the MEV.

From an organizational standpoint, the ESJQ can easily be administered to determine an employee's ethical framework, which would be beneficial in an assortment of workplace contexts. The ESJQ results would allow firms to recognize the differing ethical frameworks possessed by their employees and potentially incorporate these differences into different types of training. For example, organizational ethics training for highly consequentialist employees could focus on the repercussions (outcomes) of engaging in unethical behavior while formalist training could concentrate on the existing rules of the organization and their importance to invoke duty. This would acknowledge the fact that formalists and consequentialist speak different 'languages' when it comes to their judgements, yet each of these training approaches would garner the same result, increasing ethical behavior. Additionally, firms which are concerned about perceptions of ethical misconduct may select for highly formalist employees, since these individuals would be less prone to find acceptable benefitting from deceptive or questionable but legal activities.

It is likely that other personality traits will interact with consequentialism and formalism to influence decision making and behaviors. For example, consequentialism in a Machiavellian would manifest very differently than consequentialism in an altruist. We leave this as an opportunity for future research. Additionally, future research using the ESJQ can help us understand the impacts that formalism and consequentialism have on ethical behavior in organizations as well as their role in different cultural contexts around the world.

Conclusion

The principle aim of this work was to establish a new scale to measure consequentialism and formalism; the primary foundations of how individuals' ethically reason. These six studies, taken as a whole, provide a substantial argument for the use of the ESJQ as a way to independently measure consequentialism and formalism and they point to the predictive abilities of the ESJQ across a variety of measures and contexts.

Compliance with Ethical Standards

Conflict of interest Authors A, Author B, and Author C declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

Alder, G. S., Schminke, M., & Noel, T. W. (2007). The impact of individual ethics on reactions to potentially invasive HR practices. *Journal of Business Ethics, 75*(2), 201–214.

Banerjee, K., Huebner, B., & Hauser, M. (2010). Intuitive moral judgments are robust across variation in gender, education, politics and religion: A large-scale web-based study. *Journal of Cognition and Culture, 10*(3), 253–281.

Barak-Corren, N., & Bazerman, M. H. (2017). Is saving lives your task or God's? Religiosity, belief in god, and moral judgment. *Judgment and Decision Making, 12*(3), 280–296.

Baron, J. (2017). Utilitarian vs. deontological reasoning (p. 137). Abingdon: Taylor and Francis.

Brady, F. N. (1985). A Janus-Headed model of ethical theory: Looking two ways at Business/Society issues. *The Academy of Management Review, 10*(3), 568–576.

Brady, F. N. (1990). *Ethical managing: Rules and results*. Basingstoke: Macmillan Publishing Company.

Brady, F. N., & Wheeler, G. E. (1996). An empirical study of ethical predispositions. *Journal of Business Ethics, 15*(9), 927–940.

Burton, B., Dunn, C., & Goldsby, M. (2006). Moral pluralism in business ethics education: It is about time. *Journal of Management Education, 30*(1), 90–105.

Casali, Gian Luca (2011). Developing a multidimensional scale for ethical decision making. *Journal of Business Ethics, 104*(4), 485–497.

Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality & Social Psychology, 104*(2), 216–234.

Decety, J., Cowell, J. M., Lee, K., Mahasneh, R., Malcolm-Smith, S., Selcuk, B., & Zhou, X. (2015). The negative association between religiousness and children's altruism across the world. *Current Biology, 25*(22), 2951–2955.

Duke, A. A., & Bègue, L. (2015). The drunk utilitarian: Blood alcohol concentration predicts utilitarian responses in moral dilemmas. *Cognition, 134*, 121–127.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, *18*, 39–50.

Gefen, D., & Straub, D. (2005). A practical guide to factorial validity using PLS-Graph: Tutorial and annotated example. *Communications of the Association for Information systems, 16*(1), 5.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition, 107*(3), 1144–1154.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*(5537), 2105–2108.

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling. [White paper]. Retrieved from http://www.afhayes.com/public/process2012.pdf.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135.

Huff, C., & Tingley, D. (2015). "Who are these people?" Evaluating the demographic charasteristics and political preferences of MTurk participants. In *Research and Politics* (pp. 1–12). Abingdon: Routledge.

Hütter, M., Friesdorf, R., Armstrong, J. B., Conway, P., & Gawronski, B. (2016). Understanding responses to moral dilemmas: Deontological inclinations, utilitarian inclinations, and general action tendencies. The social psychology of morality (pp. 91–110). Abingdon: Routledge.

Koenig, H. G., & Büssing, A. (2010). The Duke University Religion Index (DUREL): A five-item measure for use in epidemological studies. *Religions, 1*(1), 78.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature, 446*(7138), 908.

Kohlberg, L. (1979). *The meaning and measurement of moral development*. Worcester, MA: Clark University Press.

Kumar, K., Borycki, C., Nonis, S. A., & Yauger, C. (1991). The strategic decision framework: Effect on students' business ethics. *Journal of Education for Business, 67*(2), 74–79.

Kumar, K., & Strandholm, K. (2002). American business education—effect on the ethical orientation of foreign students. *Journal of Education for Business, 77*(6), 345–350.

Letwin, C., Wo, D., Folger, R., Rice, D., Taylor, R., Richard, B., & Taylor, S. (2016). The "Right" and the "Good" in Ethical Leadership: Implications for Supervisors' Performance and Promotability Evaluations. *Journal of Business Ethics, 137*(4), 743–755.

Love, E., Staton, M., & Rotman, J. D. (2015). Loyalty as a matter of principle: the influence of standards of judgment on customer loyalty. Marketing Letters, 1–14.

McMahon, J. M., & Harvey, R. J. (2007). Psychometric properties of the Reidenbach–Robin multidimensional ethics scale. *Journal of Business Ethics, 72*(1), 27–39.

Middleton, K. L., & Byus, K. (2004). The relationship between precharismatic leader emergence and ethical attitudes and behaviors: Hope for the future? *Journal of Applied Management and Entrepreneurship, 9*(4), 86.

Mill, J. S. (1863). *Utilitarianism*. London: Parker, Son and Bourn.

Mudrack, P. E., & Mason, E. S. (2017). Utilitarian traits and the Janus-Headed model: Origins, meaning, and interpretation. *Journal of Business Ethics*. https://doi.org/10.1007/s10551-017-3592-5.

Nozick, R. (1981). *Philosophical explanations*. Boston, MA: Harvard University Press.

O'Shaughnessy, N. (2002). Toward an ethical framework for political marketing. *Psychology and Marketing, 19*(12), 1079–1094.

Paxton, J. M., Ungar, L., & Greene, J. D. (2012). Reflection and reasoning in moral judgment. *Cognitive Science, 36*(1), 163–177.

Piazza, J. (2012). "If You Love Me Keep My Commandments": Religiosity increases preference for rule-based moral arguments. *International Journal for the Psychology of Religion, 22*(4), 285–302.

Pearsall, M. J., & Ellis, A. P. J. (2011). Thick as thieves: The effects of ethical orientation and psychological safety on unethical team behavior. *Journal of Applied Psychology, 96*(2), 401–411.

Piazza, J., & Landy, J. F. (2013). "Lean not on your own understanding": Belief that morality is founded on divine authority and non-utilitarian moral judgments. *Judgment and Decision Making, 8*(6), 639.

Piazza, J., & Sousa, P. (2013). Religiosity, political orientation, and consequentialist moral thinking. *Social Psychological and Personality Science, 5*(3), 334–342.

Rajczi, A. (2016). On the incoherence objection to rule-utilitarianism. *Ethical Theory and Moral Practice, 19*(4), 857–876.

Reidenbach, R. E., & Robin, D. P. (1988). Some initial steps toward improving the measurement of ethical evaluations of marketing activities. *Journal of Business Ethics, 7*(11), 871–879.

Reidenbach, R. E., & Robin, D. P. (1990). Toward the development of a multidimensional scale for improving evaluations of business ethics. *Journal of business ethics, 9*(8), 639–653.

Reynolds, S. J. (2006). Moral awareness and ethical predispositions: Investigating the role of individual differences in the recognition of moral issues. *Journal of Applied Psychology, 91*(1), 233–243.

Reynolds, S. J. (2008). Moral attentiveness: Who pays attention to the moral aspects of life? *Journal of Applied Psychology, 93*(5), 1027–1041.

Reynolds, S. J., & Ceranic, T. L. (2007). The effects of moral judgment and moral identity on moral behavior: An empirical examination of the moral individual. *Journal of Applied Psychology, 92*(6), 1610–1624.

Rotman, J. D., Khamitov, M., & Connors, S. (2018). Lie, Cheat, and Steal: How Harmful Brands Motivate Consumers to Act Unethically. *Journal of Consumer Psychology*. https://doi.org/10.1002/jcpy.1002.

Sashkin, M., Rosenbach, W. E., & Sashkin, M. G. (1997) "Development of the power need and its expression in leadership and management with a focus on leader-follower relations." *Leadership as legacy: Proceedings of the twelfth scientific meeting of the AK Rice Institute*. AK Rice Institute, Jupiter, FL.

Schminke, M. (1997). Gender differences in ethical frameworks and evaluation of others' choices in ethical dilemmas. *Journal of Business Ethics, 16*(1), 55–65.

Schminke, M. (2001). Considering the business in business ethics: An exploratory study of the influence of organizational size and structure on individual ethical predispositions. *Journal of Business Ethics, 30*(4), 375–390.

Schminke, M., Ambrose, M. L., & Noel, T. W. (1997). The effect of ethical frameworks on perceptions of organizational justice. *Academy of Management Journal, 40*(5), 1190–1207.

Sinnott-Armstrong, W. (2015). Consequentialism. *The Stanford Encyclopedia of Philosophy*. Winter 2015. Retrieved from https://plato.stanford.edu/archives/win2015/entries/consequentialism/.

Smith, C. G., Šumilo, E., & Pauls Karnups, V. (2009). Moral judgment: A comparison of Latvian and US business persons. *Baltic Journal of Management, 4*(2), 188–205.

Vitell, S. J., & Muncy, J. (2005). The Muncy–Vitell consumer ethics scale: A modification and application. *Journal of Business Ethics, 62*(3), 267–275.

Walsh, R. T. G. (2015). Introduction to Ethics in Psychology: Historical and Philosophical Grounding. *Journal of Theoretical and Philosophical Psychology, 35*(2), 69–77.