Using Entropy-Based Information Theory to Evaluate Survey Research

Daniel Friesner
North Dakota State University

Faith Valente
North Idaho College

Carl S. Bozman
Gonzaga University

An assumption underlying most survey-based research studies is that respondents provide equally informative responses. While researchers regularly examine the validity and reliability of survey data, few assess whether the quantity of information collected is consistent across respondents. This manuscript uses information theory to evaluate whether certain types of respondents are more (or less) likely to give equally informative responses. A simple method to evaluate the quantity of information inherent in survey items and sampling designs is presented. We apply this technique to a survey conducted for a large, amateur sporting event and find that the survey design was generally appropriate.

INTRODUCTION

Over the past several decades, a significant portion of the marketing research literature has focused on designing, administering, and assessing survey-based data collection and analysis methods (Hsia, 1988; Dillman, 2000; Berger, 2013). One dimension of this literature has focused on enhancing survey design techniques. That is, by carefully crafting survey related items, and by strategically choosing the most appropriate measures, respondents are likely to give more informative responses. Another element of this literature has evaluated survey administration within the context of research design. By altering who receives an opportunity to participate in a survey (whether by sampling technique or experimental procedure), researchers can collect a statistically meaningful approximation of the attitudes, values, and beliefs of a population using fewer resources. A final component of this literature assesses the outcomes associated with effective survey design and administration. If a survey words items appropriately, provides an appropriate response scale, and is administered to a representative sample of respondents (who are properly motivated to complete and return the survey), the data collected can be analyzed and reported utilizing appropriate statistical methods. The data are said to be “valid” and “reliable” reflections of the underlying population. Any descriptive or inferential statistical analyses made using the data (if the method of analysis makes optimal use of the data) are deemed to be “unbiased” and “efficient”. Indeed, the validity and reliability of the data, and the resulting unbiasedness and efficiency of statistical analyses, are the primary criteria upon which the usefulness of survey-based research projects are usually based (Cottrell & McKenzie, 2005; Smith & Albaum, 2005).

Although many contributions to each portion of the research method literature have been made, Eisend (2015) notes that the aggregate improvement in research methods are insignificant when compared to knowledge acquisition in marketing related subject areas. For example, the advertising or integrated communications field, as a whole, has matured more rapidly than the development of the research methods which are used to evaluate theories produced in this area of study. Advertising inquiries have become more specialized and incremental in nature as a consequence of our more advanced understanding of the integrated communications nexus. Research methods are needed which better accommodate the correspondingly broader range of communication phenomena as well as improve discrimination among communication effects. Otherwise, the empirical methods employed by survey researchers may prove less than adequate and lead to confounding conclusions. As an illustrative example, consider the concept of empathy; an important communication ability people use to implicitly understand and relate to the world around them (Hoffman, 2000). This automatic or reflective identification with another person's perceived thoughts, feelings, and behaviors is a perceptive state that can be induced by communication stimuli. Lijiang (2010) describes how empathetic messages have an effect on knowledge acquisition and persuasion. Several other researchers have shown that empathy has no such effect on consumer reactions to communication stimuli (Massi Lindsey, Yun, & Hill, 2007; Merolla, Zhang, & Sun, 2013). One plausible explanation for the former inconsistent findings, with regard to the effects of empathetic communications, may be different quantities of information provided by alternative groups of research participants.

Consider the potential confounding effects of community boundedness on the empathetic perceptions of people. Community boundedness describes the differential extent to which an issue affects groups of people in a population and which, as a result, may affect the quantity of information those respondents provide (Rucinski, 2004). Community boundedness exerts a direct influence on knowledge acquisition. More specifically, the previous author documents how people of higher social-economic status knew less about a medical insurance program targeted at the poor than did people of a lower social-economic status regardless of personal relevance. Bas and Grabe (2015) also describe how differences in the emotional response among people, due to empathetic stimuli, results in smaller knowledge discrepancies between higher and lower educated groups.

Yet a crucial assumption underlying quantitative survey research analyses is that a specific respondent (whether overall or within a given aspect of a research design, such as a sampling stratum) will provide the same quantity of information in his or her responses as every other respondent in the same aspect of the research design. There is little value in examining the *quality* of the information contained in a data set (i.e., the data's reliability and validity, and the unbiasedness and efficiency of the statistical results), if this assumption is violated and the *quantity* of information provided by respondents is minimal or varies substantially across respondents (Friesner, Khayum, & Schibik, 2013). Any research procedure that could be used to assess the veracity of the equality in information assumption across groups of empathetic respondents, or any other respondent groups, could be a significant contribution to research methodology as applied within any sub-field of marketing, including integrated communications.

To illustrate the value of this distinction, consider the common research analogy of an archer shooting arrows at a target, where each arrow represents a respondent's completed survey responses. The concepts of validity and (in its application in statistics) unbiasedness are concerned with whether the arrows hit (or miss) the target, while reliability and (again, in its use in statistical analysis) efficiency pertains to the pattern with which the arrows cluster on the target. If the data are valid and the statistical results are unbiased, the vast majority of the arrows hit the target near its center. If the data are reliable (and efficient), the arrows are grouped very closely together, hopefully (if valid and unbiased) near the center of the target. In this analogy, the *quantity* of information refers to the size and shape of the target itself. If the target is very large (i.e., respondents provide large quantities of information) it is easier to hit the target and cluster arrows near the center of the target. If the target is small (little information is provided by respondents) it is more difficult, and perhaps impossible, to either hit the target or to gauge whether the arrows are clustering together.

In most research applications, any respondent characteristics that may impact the quantity of information in survey responses, and which are not explicitly controlled for in the research design, are assumed to be constant across respondents (Cottrell & McKenzie, 2005; Dowling, 2015). This assumption may often be inappropriate. For example, in pharmaceutical marketing studies, one must assume that all individuals who receive pharmacy services, or who are candidates to be prescribed a specific medication, have the same levels of health literacy (Mackert, Guadagno, Mabry, & Chilek, 2013). They must not only be equally informed about the nature of the medication, but must read and interpret survey items, which collect patient attitudes, values, beliefs, and preferences, in a consistent manner. When this assumption does not hold, it would be incorrect to treat each respondent's completed surveys in an equal manner and give them equal weight when statistically analyzing those responses. In such cases, the research design should be adapted (usually through stratification) to ensure that more informative responses receive different weight in subsequent statistical analyses (Dahl & Osteras, 2010).

This raises two interesting questions: first, is it possible to evaluate the quantity of information that respondents provide? And if so, is it possible to empirically identify those respondents who give more (or less) informative responses? This paper argues that the answer to these questions, from an *ex post* basis, can be "yes". More specifically, we demonstrate how information theoretic techniques can be applied to evaluate whether empathetic respondents, when evaluated over an array of related survey items, are likely to give more (or less) informative responses (Dahl & Osteras, 2010; Schibik, Khayum, & Friesner, 2012). In doing so, we provide researchers with a simple means to evaluate the quantity of information inherent in their survey items and sampling designs. As an empirical illustration, these methods are used to retrospectively evaluate the quantity of information collected from a survey-based research study conducted at a large, amateur sporting event. Large scale survey research projects (for example, clinical drug trials) which utilize pilot studies prior to full scale implementation could also use these tools prospectively (by analyzing the pilot study data and adjusting the full scale study accordingly) to maximize the quantity of information collected from respondents (Masri et al. 2012; Fiedler & Bebber, 2013; Fiedler, Bebber, & Oetjen, 2013).

The remainder of the paper proceeds in several steps. In the next section, we describe how measures of entropy, which are drawn from the information theory literature, can be used to capture the quantity of information in a series of related survey items. This leads to the creation of a testable hypothesis about the quantity of available information, around which an empirical model can be constructed to evaluate this hypothesis. The third section describes the data we use to illustrate our methodology and test our hypothesis, which come from an established marketing research study that has been implemented repeatedly at a well-known, annual, amateur basketball tournament (Bozman, Kurpis, & Fry, 2010; Kurpis, Bozman, & Kahle, 2010). Results are presented and interpreted in the fourth section. We conclude the paper by summarizing our findings, discussing their implications to the marketing research literature, and suggesting some future areas of inquiry related to information entropy.

ENTROPY-BASED INFORMATION MEASURES

The use of entropy to measure the quantity of information contained in a survey rests upon several assumptions. The first, and most important, assumption is a prior expectation of ignorance on the part of the researcher (Jaynes, 1957, 1982). Within the context of survey design, this implies that a survey is designed and administered such that the researcher has no prior expectations concerning the distribution of responses for a given survey item or scale (collected over a set of individuals), or a given individual (collected over a set of survey items or scales). That is, the ignorance assumption requires that the survey is designed to minimize the likelihood of leniency, common method variance and/or framing biases (among other design issues) which reduce the sensitivity of the survey item(s) or scale(s) being analyzed (Smith & Albaum, 2005). Under this prior expectation, the expected distribution of responses is uniform.

The concept of entropy can be effectively applied to survey items (or scales) with mutually exclusive and collectively exhaustive, multiple choice responses (Cox, 1980). In a given survey item, respondents typically choose from a discrete set of k = 1,…,K possible responses.[1] For each possible response k, it is possible to calculate the proportion of responses (p_k) that fall into that category for a given sample of data. Consistent with most statistical principles, information theory requires that the proportions be proper; that is: $0 \leq p_k \leq 1$ and $\sum_{k=1}^{K} p_k = 1$. Shannon (1948), Finn and Roberts (1984), Golan (2006), and Golan, Judge, and Miller (1996), among others, have demonstrated that the entropy, or quantity of information in the system of responses, can be characterized as:

$$H(p) = -\sum_{k=1}^{K} p_k log_2(p_k) \tag{1}$$

where H indicates the entropy measure and $log_2()$ represents the base two logarithm. It is also customary to assume that $p_k log_2(p_k) = 0$ when $p_k = 0$ (Golan et al. 1996). Under this formulation, entropy is maximized, when $p_k = 1/K$, for every k. In other words, entropy is maximized when the researcher's assumption of ignorance is appropriate and the proportions contain as little information as possible, i.e. no deviation from the expected distribution. Any empirical realization of p_ks that deviate from maximum entropy also deviate from these prior expectations. By focusing on the distribution of responses, the entropy measure simultaneously encompasses measures of central tendency (which measure specific types of validity) as well as measures of variability (which are used to assess reliability) in the data.

Dahl and Osteras (2010) argue that, in relative terms, the total amount of information that can be gleaned from this variable is given by the ratio of the actual, estimated entropy, divided by the maximum possible entropy value. This ratio can be used in survey design as a guide to choosing the number of points (K) in the response scale. As an example, the authors considered the Norwegian Functional Assessment Scale which is available in two versions, one of which employs a 4 point response scale for each of the 39 items in the survey, while the other employs a 5 point scale for the same 39 items. The authors found that the average percentage of information contained in the 4 point scale was 34.5 percent, while the 5 point scale extracted 40.1 percent of maximum available information in the data. Hence the 5 point scale is preferred to the 4 point scale on the grounds that (all else constant) an optimally designed response scale should (a priori) provide a distribution of responses that is as close to uniformly distributed (and is as close to maximum entropy) as possible.

Researchers typically ask a sample of respondents to address multiple items in a survey. Let l=1,…,L denote the number of survey items (or the same survey item over time if the survey is administered repeatedly) and let i=1,…,n denote the respondents in the sample. For simplicity, we further assume that each of the n respondents completes the same set of L survey items, and that each of the L survey items uses the same K response scales.[2] This provides two possible methods of analyzing the information content in a given survey. One method is to aggregate over individuals and examine the distribution of p_ks across a series of survey items, in which case $p_{kl} = \frac{\sum_{i=1}^{n} D_{kli}}{n_l}$, for each k = 1,…,K, where D is a binary (or dummy variable) indicator that gives a value of 1 if respondent i gave response k for survey item l, and a value of zero otherwise. Entropy then is calculated as a variable over l = 1,…,L units of observation (time periods) and subsequently recorded as a variable, which can be analyzed statistically:

$$H_l(p) = -\sum_{k=1}^{K} p_{kl} log_2(p_{kl}) \tag{1b}$$

Friesner et al. (2013), for example, use a series panel of (repeatedly administered) business outlook surveys conducted by the New York Federal Reserve Bank to characterize the quantity of information contained in specific types of business outlook surveys.

Alternatively, if one is interested in determining which survey respondents provide greater or lesser quantities of information across individual respondents, it is necessary to examine a single administration of a survey and aggregate responses for each individual over a series of L related questions. This leads to a slightly different characterization of the information entropy formulation. Define: $p_{ik} = \frac{\sum_{l=1}^{L} D_{kli}}{L_i}$, for each k = 1,…,K, where D is defined previously and L_i is the total number of survey items answered by respondent i.[3] Entropy then is calculated as a variable over the n observations in the sample:

$$H_i(p) = -\sum_{k=1}^{K} p_{ki} log_2(p_{ki}) \tag{1c}$$

Consistent with our previous discussion, equation (1c) can be expressed as an absolute measure, or divided by the maximum entropy attainable over the L possible survey items and subsequently expressed as a proportion of total entropy in the system.

Equation (1c) provides the basis for our testable hypothesis. Our prior assumption is that of ignorance; each respondent who completes the survey is equally likely to give any possible response to a given survey item. Thus, over L survey items, each respondent is equally likely to select any of the K possible responses for any of the L survey items. If, over the L items in a survey, respondents provide an equal quantity of information in their responses, then the entropy calculation provided in (1c) should not vary systematically across individuals, or across groups of individuals who share similar socio-demographic characteristics. The null and alternative hypotheses in this study can therefore be expressed as follows:

H_0: No mean differences in information entropy exist across groups of respondents
H_A: Mean differences in information entropy exist across groups of respondents

Concomitantly, if respondents systematically give non-uniform responses across the L items in the survey based on one or more socio-demographic factors, evidence exists to suggest that different groups of respondents give more (or less) information when completing those survey items. For any single social or demographic characteristic, one can calculate the entropy measure in (1c) and apply analysis of variance (or its nonparametric analog, the Kruskal-Wallis test) to reject or fail to reject the null hypothesis (Smith & Albaum, 2005; Slack & Baidoo, 2015). Alternatively, regression analysis can be used to assess the null hypothesis when an individual respondent may belong to a variety of different demographic groups and/or express a variety of different social characteristics (Smith & Albaum, 2005; Slack & Baidoo, 2015).

Assume that each survey respondent was randomly selected to participate in the survey and that no selection or other response biases exist in the survey's administration. Then define ε_i as a random error term (over i =1,…,n observations), define X^q as one of Q respondent social characteristics or demographic variables, define α as an estimated intercept and β^q as one of q = 1,..,Q slope parameters to be estimated. The regression equation of interest can consequently be stated as linear in parameters with the following form:

$$H_i(p) = \alpha + \sum_{q=1}^{Q} \beta^q X_i^q + \varepsilon_i \tag{2}$$

Under the null hypothesis, each parameter estimated (whether tested individually or jointly) should be statistically no different from zero. That is, for individual characteristics:

$$\begin{aligned} &H_0\text{: } \beta^q = 0 \text{ for each } q = 1,\ldots, Q \\ &H_A\text{: } \beta^q \neq 0 \text{ for each } q = 1,\ldots, Q \end{aligned} \tag{3}$$

For joint tests:

$$\begin{aligned} &H_0\text{: } \beta^1 = \beta^2 = \beta^Q = 0 \\ &H_A\text{: Not } H_0 \end{aligned} \tag{4}$$

Having specified our basic statistical analysis, it is necessary to address some ancillary issues that pertain to estimating equation (2). First, traditional entropy measures are unbounded on the lower end, but have a fixed upper bound by construction, which varies based on the number of survey items (L) and the number of possible responses (K) in each item. To address this issue, we normalize our entropy measure by dividing each entropy value by its maximum. This yields a normalized entropy measure, which is consistently bounded on the closed unit interval, and (given an appropriate, randomly identified survey administration) facilitates estimation using a two-sided Tobit model (Golan et al., 1996; Greene, 2000). Hypothesis tests on individual parameters can be conducted using standard t-tests, while joint tests of parameters can be conducted using a (likelihood ratio) chi-square test. In all cases, we use a 5 percent level of statistical significance, although significance at the 10 percent level is also noted. As in most regression analysis, it is common to identify certain types of demographic characteristics (for example, respondent gender) using a set of F mutually exclusive and collectively exhaustive dummy variables. To avoid perfect multicollinearity, we drop one of these variables and include F-1 of these variables in the regression (Greene, 2000). All resulting parameter estimates for the included dummy variables are interpreted relative to the omitted category.

Having specified our null hypothesis and model of analysis, it is important to note several practical considerations when applying this methodology. First, if a survey is designed and administered appropriately, one should expect that no statistical differences in the quantity of information exist across groups of respondents. In other words, *a lack of statistical significance across (groups of) respondents is evidence (but not proof) that the survey was designed and administered appropriately.* Rejection of the null hypothesis (whether overall or for specific subsets of respondents) may indicate the need to adjust the survey's design and/or administration to better capture the information being provided by respondents.

Second, if one rejects the null hypothesis for a group of respondents, it implies that the group identified by the significant estimate(s) is giving a fundamentally different set of responses, as measured by information entropy, than other groups of respondents. Under the strictest interpretation of entropy, any effect that moves a respondent's entropy away from the maximum value (i.e., results that are non-uniform, or clustered towards either end of a response scale) leads to the collection of less information than effects that move responses closer to the maximum entropy value. However, and unlike the application of entropy provided by Dahl and Osteras (2010), this does *not* mean that these groups are less interesting to the researcher to study. Rather, any such findings imply that these respondents simply answered the survey in a fundamentally *different* manner than others who completed the survey and could result in misleading interpretations should group responses be aggregated. How researchers address these groups (through stratified sampling, by altering the response scales, changing the wording of the survey items, etc.) must be determined by the researcher based on the population being studied and the phenomena being investigated.

DATA

The data we evaluate are drawn from a survey administered at the annual Hoopfest basketball tournament: the largest 3-on-3 amateur basketball tournament in the world (Schnell, 2014). One quarter of a million people were estimated to attend and participate in or watch this three-day basketball event. The data used in this study come from a survey administration during the 2013 Hoopfest tournament, and its overall design and administration, based on previous iterations of the survey, have been discussed elsewhere in the literature (Bozman et al., 2010; Kurpis et al., 2010). Survey responses were anonymous, and the (anonymous) survey data are freely available on the Hoopfest website: http://www.spokanehoopfest.net/organization/Pages/history.aspx. Because the data are anonymous, analyzed as secondary data and lie within the public domain, our analysis is not considered human subjects research and was not subject to institutional research board approval.

The primary survey elements of interest in this study are questions 7a through 7i, which together comprise a nine-item values scale. These nine scale items have previously been shown to demonstrate evidence of both nomological and predictive validity (Bearden & Netemeyer, 1999). More specifically, these nine value scale items highlight how social affiliation values, personally oriented values, and external stimulation values influence what people do as they go about their daily lives (Kahle, 1983; Homer & Kahle, 1988).

The nine scale items require respondents to rate how much they value specific activities and how much they want to experience those activities or accomplishments in their daily lives. The activities or accomplishments include: having a sense of belonging, experiencing excitement, having a warm relationship with others, experiencing self-fulfillment, being respected by others, experiencing fun and enjoyment in life, having security, having self-respect, and experiencing a sense of accomplishment. Individuals are asked to rate their responses on a 1 to 9 scale, with 1 being "Very Unimportant" and 9 being "Very Important".

Consistent with equation (1c), we calculate a single entropy measure for each individual by identifying the proportion of an individual's responses over the nine survey items ($L = 9$) that a person gave each of the $K = 9$ possible scale responses. These survey items are generally applicable to the human condition, worded simply and clearly, and the response scales are sufficiently clear and broad in anchoring to meet the assumptions of the empirical model. In other words, the questions should apply equally and be equally interpreted by respondents. The responses should be sufficient to accurately and precisely assess information entropy in such a way as to support or reject the maintained assumption of ignorance concerning the actual, empirical distribution of responses.

Examination of the survey suggests that several possible remaining survey items can be used to classify or categorize respondents. For example, the survey asks respondents to reveal their gender, age (in years), zip code of residence, whether they are attending Hoopfest because they are participating as a player, spectator, volunteer, or for some other reason. Respondents are also asked to report the number of times they have attended a Hoopfest tournament, their level of satisfaction with the tournament, whether each respondent plans to attend future Hoopfest tournaments, the respondent's perception of the importance of the Hoopfest tournament to the local community, the respondent's perceptions regarding her/his ability to get caught up in others' feelings, and the respondent's perceptions regarding his or her empathy towards others. The respondent's age and number of times attending Hoopfest were recorded as quantitative variables. The remaining variables (which are themselves discrete or not directly interpretable in a quantitative and marginal fashion) were discretized into a series of mutually exclusive and collectively exhaustive dummy variables. In most cases a dummy variable was created for each possible response to a given question. However, to prevent multicollinearity in the regression analysis, some of these discrete variables were also combined into more aggregated groups of responses. For example, in cases where a five point Likert scale is used (strongly disagree, disagree, neither agree nor disagree, agree, and strongly agree), the results were aggregated into three categories by combining the strongly agree and agree responses together, as well as combining the strongly disagree and disagree responses.

The full sample contained 418 responses. After eliminating individuals who failed to provide a complete set of responses, we are left with a working sample of 314 responses, which represents a 75.1 percent survey completion rate.

RESULTS

Table 1 contains the names, definitions and descriptive statistics for each of the variables used in the analysis. Given that there are nine possible responses in the scale, the maximum possible entropy available in the system is 3.170. The mean level of entropy in the sample is 1.005, which implies that slightly more than thirty percent (31.7 percent) of the maximum possible entropy is captured in the responses. This number is consistent with the percentage of entropy captured in the Dahl and Osteras study (2010).

Approximately half (47.5 percent) of respondents were female. The mean age of respondents was 32.65 years of age, and only 6.3 percent of respondents were minors. Slightly more than half (51.3 percent) of the respondents attended Hoopfest in order to participate as a player, while approximately

30.6 percent participated as a viewer, 14.6 percent participated as a volunteer, and the remainder (3.5 percent) participated in other ways. At the mean, the typical respondent had previously attended 6.52 previous Hoopfest events. Approximately 78.9 percent of participants lived outside of the host city (Spokane, WA), with the remaining 21.1 percent living in various areas within Spokane. Within the sample, 98.4 percent of respondents indicated they were satisfied with the Hoopfest event, and 89.8 percent indicated the intent to attend future Hoopfest events. Approximately 99.4 percent of respondents reported a belief that Hoopfest was good for the local community. With regard to empathy, 65.6 percent of respondents reported that they have a degree of empathy and are able to understand how people feel before being told, while 25.8 percent provided a neutral response and 8.6 percent reported that they disagreed or strongly disagreed with the notion that they were empathetic. Concomitantly, 26.4 percent of respondents agreed or strongly agreed that they get caught up in the feelings of others, while 39.8 percent disagreed or strongly disagreed with this statement. The remaining 33.8 percent of respondents neither agreed nor disagreed with this statement.

TABLE 1
DESCRIPTIVE STATISTICS

Variable	Description	Mean	Std. Dev.
ENTROPY	Entropy Calculation	1.005	0.753
NENTROPY	Normalized (Proportional) Entropy Calculation	0.317	0.238
PLAY	Binary variable identifying participants who play in the Hoopfest tournament	0.513	
WATCH	Binary variable identifying participants who primarily watch the Hoopfest tournament	0.306	
VOL	Binary variable identifying participants who primarily volunteer at the Hoopfest tournament	0.146	
OTHER	Binary variable identifying participants who attend the Hoopfest tournament for other reasons	0.035	
SATIS	Binary variable identifying participants who are satisfied (slightly, moderately or very) with the Hoopfest tournament	0.984	
DISSATIS	Binary variable identifying participants who are dissatisfied (slightly, moderately or very) with the Hoopfest tournament	0.016	
NVSATIS	Binary variable identifying participants who are not very satisfied with the Hoopfest tournament	0.309	
NOATT	The total number of times participants report attending the Hoopfest tournament	6.519	6.019
ATTEND	Binary variable identifying participants who definitely or probably will attend the Hoopfest tournament in 2014	0.898	
UATTEND	Binary variable identifying participants who may or may not attend the Hoopfest tournament in 2014	0.083	
NATTEND	Binary variable identifying participants who definitely or probably will not attend the Hoopfest tournament in 2014	0.019	
IMPCOMM	Binary variable identifying participants who believe that the Hoopfest tournament is (slightly, moderately or very) important to the Spokane community	0.994	

UCOMM	Binary variable identifying participants who believe that the Hoopfest tournament is (slightly, moderately or very) unimportant to the Spokane community	0.006	
NIMPCOMM	Binary variable identifying participants who believe that the Hoopfest tournament is not very important to the Spokane community	0.146	
DFEEL	Binary variable identifying participants who disagree or strongly disagree that they get caught up in other people's feelings easily	0.398	
UFEEL	Binary variable identifying participants who neither agree nor disagree that they get caught up in other people's feelings easily	0.338	
AFEEL	Binary variable identifying participants who agree or strongly agree that they get caught up in other people's feelings easily	0.264	
DEMPATHY	Binary variable identifying participants who disagree or strongly disagree that they can often understand how people are feeling before they are told	0.086	
UEMPATHY	Binary variable identifying participants who neither agree nor disagree that they can often understand how people are feeling before they are told	0.258	
AEMPATHY	Binary variable identifying participants who agree or strongly agree that they can often understand how people are feeling before they are told	0.656	
AGE	Respondent age in years	32.653	16.207
AGEU18	Binary variable identifying respondents under age 18	0.194	
AGE1829	Binary variable identifying respondents aged 18-29	0.299	
AGE3039	Binary variable identifying respondents aged 30-39	0.191	
AGE4049	Binary variable identifying respondents aged 40-49	0.156	
AGE50P	Binary variable identifying respondents under aged 50 and older	0.159	
ZSPOKANE	Binary Variable identifying Spokane residents by zip code	0.213	
Z99223	Binary Variable identifying Spokane residents by zip code 99223	0.048	
Z99206	Binary Variable identifying Spokane residents by zip code 99206	0.048	
Z99203	Binary Variable identifying Spokane residents by zip code 99203	0.022	
Z99204	Binary Variable identifying Spokane residents by zip code 99204	0.003	
Z99202	Binary Variable identifying Spokane residents by zip code 99202	0.016	
Z99201	Binary Variable identifying Spokane residents by zip code 99201	0.019	
Z99207	Binary Variable identifying Spokane residents by zip code 99207	0.022	

Z99205	Binary Variable identifying Spokane residents by zip code 99205	0.035
ZOTHER	Binary Variable identifying non-Spokane residents by zip code	0.787
FEMALE	Binary Variable identifying female respondents	0.475
Number of Observations		314

Table 2 contains Analysis of Variance (ANOVA) and Kruskal-Wallis tests to determine whether the distribution of entropy varies systematically by respondent characteristics. No statistically significant differences exist in the quantity of information across respondents based on their ability to get caught up in others' feelings (ANOVA prob. = 0.178; Kruskal-Wallis prob. = 0.276), perceived empathy (ANOVA prob. = 0.444; Kruskal-Wallis prob. = 0.458), reason for attending Hoopfest (ANOVA prob. = 0.831; Kruskal-Wallis prob. = 0.662), intention to attend future Hoopfest events (ANOVA prob. = 0.231; Kruskal-Wallis prob. = 0.130), by age category (ANOVA prob. = 0.913; Kruskal-Wallis prob. = 0.936), area of residence (ANOVA prob. = 0.181; Kruskal-Wallis prob. = 0.169) or by gender (ANOVA prob. = 0.090; Kruskal-Wallis prob. = 0.051).

Statistically significant differences in information entropy did exist across respondents based on their levels of satisfaction with the event (ANOVA prob. = 0.014; Kruskal-Wallis prob. = 0.016), with the highest entropy values reported among those who were slightly satisfied (mean entropy = 0.487) and moderately satisfied (mean entropy = 0.344) compared to those who were dissatisfied (mean entropy = 0.206) and very satisfied (mean entropy = 0.299). Statistically significant differences in entropy also existed across respondents based on their perceptions concerning the importance of the Hoopfest event to contribute to the local community (ANOVA prob. = 0.008; Kruskal-Wallis prob. = 0.009), with lower entropy values reported among those who felt the event was very important (mean entropy = 0.299) compared to all other responses. The entropy of responses rose consistently and dramatically with the negativity of those perceptions. For example, those who felt the Hoopfest event was moderately unimportant provided a mean entropy of 0.763. Cumulatively, these results suggest that the null hypothesis can be rejected, especially among those respondents who differ in their perceptions of the community impact of the Hoopfest event and who are differentially satisfied with the event as a whole. It is also important to note that these results, while statistically significant, are also the variables for which respondents gave very skewed responses. For example, only five individuals believed that the Hoopfest tournament was only slightly important or unimportant to the local community. Moreover, only five respondents were dissatisfied (to any degree) with the Hoopfest tournament. Hence, it is possible that the significant differences in entropy exhibited in these two variables may be driven by extreme preferences at the tails of the distributions of these variables.

TABLE 2
MEAN COMPARISONS

Dependent Variable:			**NENTROPY**			
			ANOVA		**Kruskal-Wallis**	
Variable	**Obvns.**	**Mean**	**Statistic**	**Prob.**	**Statistic**	**Prob.**
I get caught up in other people's feelings easily						
Strongly Disagree	47	0.369	1.586	0.178	5.114	0.276
Disagree	78	0.344				
Neither Agree nor	106	0.303				

Disagree								
Agree	54	0.295						
Strongly Agree	29	0.251						
I can often understand how people are feeling even before they tell me								
Strongly Disagree	7	0.394	0.935	0.444		3.635	0.458	
Disagree	20	0.323						
Neither Agree nor Disagree	81	0.288						
Agree	138	0.340						
Strongly Agree	68	0.296						
The primary reason you are at Hoopfest								
Play in Hoopfest	161	0.315	0.293	0.831		1.590	0.662	
Watch Hoopfest	96	0.038						
Volunteer at Hoopfest	46	0.330						
Another Reason	11	0.372						
How satisfied are you with Hoopfest?								
Very Satisfied	217	0.299	3.572	0.014	**	12.586	0.006	**
Moderately Satisfied	78	0.344						
Slightly Satisfied	14	0.487						
Slightly Dissatisfied	5	0.206						
Moderately Dissatisfied	0	NA						
Very Dissatisfied	0	NA						
How likely are you to attend Hoopfest next year?								
Definitely will Attend	201	0.301	1.410	0.231		7.114	0.130	
Probably will Attend	81	0.339						
May or May not Attend	26	0.399						
Probably will not Attend	4	0.277						
Definitely will not Attend	2	0.156						
How important do you feel Hoopfest is to the Spokane								

community?								
Very Important	268	0.299	3.525	0.008	**	13.442	0.009	**
Moderately Important	42	0.408						
Slightly Important	2	0.480						
Slightly Unimportant	1	0.597						
Moderately Unimportant	1	0.763						
Very Unimportant	0	NA						
Age Categories								
Under 18	61	0.343	0.245	0.913		0.817	0.936	
18 - 29	94	0.307						
30 - 39	60	0.316						
40 - 49	49	0.309						
50 and Older	50	0.313						
Zip Code of Residence								
99206	15	0.294	1.435	0.181		11.626	0.169	
99205	11	0.246						
99207	7	0.365						
99201	6	0.127						
99202	5	0.437						
99204	1	0.763						
99203	7	0.288						
99223	15	0.384						
Another Zip Code (Non-Local Resident)	247	0.317						
Gender								
Female	149	0.293	2.887	0.090	*	3.794	0.051	*
Male	165	0.339						
**** Indicates statistical significance at the 5 percent level or better**								
*** Indicates statistical significance at the 10 percent level or better**								

Table 3 contains the results of the Tobit analysis, which examines the impact of respondent characteristics on the entropy of their responses to question 7, while controlling for all other specific group specific characteristics. As noted by the chi-square statistic's probability value (0.008), the regressors jointly explain a significant percentage of variation in the dependent variable, which indicates that when taken collectively, these respondent characteristics do impact the quantity of information respondents provided. Hence, the null hypothesis is rejected in the regression results as well. The Tobit disturbance term is also positive and statistically significant from zero (coefficient estimate = 0.288; prob.

< 0.001), which indicates the necessity of econometrically adjusting for the censoring of the distribution at zero and one.

Examining the individual slope estimates in the regression, three coefficient estimates are statistically different from zero at the five percent level or better.Those that disagree or strongly disagree that they get caught up in others' feeling are significantly likely to provide higher levels of entropy (or provide more informative responses) to the items contained in question seven of the survey (coefficient estimate = 0.081; prob. = 0.050). Those individuals who agree or strongly agree that they have empathy for others are also significantly likely to provide responses with greater quantities of information than individuals who are less likely to agree to being empathetic (coefficient estimate = 0.082; prob. = 0.048). Lastly, those individuals who do not believe that Hoopfest is important to the local community are likely to provide significantly more informative responses than those who do not have such views (coefficient estimate = 0.151; prob. = 0.003).

TABLE 3
TOBIT ANALYSIS OF THE INFORMATION ENTROPY MEASURE

Dependent Variable:		**NENTROPY**			
Regressor	**Coeff.**	**Std. Error**	**t-Statistic**	**P-value**	
Intercept	0.056	0.075	0.750	0.451	
DFEEL	0.081	0.041	1.960	0.050	**
AFEEL	0.000	0.046	0.010	0.994	
DEMPATHY	0.080	0.069	1.160	0.246	
AEMPATHY	0.082	0.041	1.980	0.048	**
NIMPCOMM	0.151	0.051	2.980	0.003	**
WATCH	0.033	0.047	0.720	0.474	
VOL	0.069	0.054	1.290	0.198	
OTHER	0.135	0.098	1.370	0.169	
NVSATIS	0.060	0.038	1.570	0.117	
NOATT	0.004	0.003	1.380	0.166	
UATTEND	0.114	0.063	1.810	0.070	*
NATTEND	-0.093	0.129	-0.720	0.472	
Z99223	0.071	0.082	0.860	0.390	
Z99206	-0.047	0.083	-0.570	0.570	
Z99203	0.001	0.114	0.010	0.995	
Z99204	0.538	0.297	1.820	0.070	*
Z99202	0.177	0.133	1.330	0.183	
Z99201	-0.267	0.137	-1.950	0.051	*
Z99207	0.035	0.116	0.300	0.763	
Z99205	-0.020	0.093	-0.210	0.834	
FEMALE	-0.037	0.038	-0.970	0.332	
AGEU18	0.063	0.056	1.130	0.258	
AGE	0.001	0.001	0.580	0.562	
Tobit Disturbance Term	0.288	0.015	19.860	<0.0001	**

Log-Likelihood Function	-124.317		
Restricted Log-Likelihood Function	-145.639		
Chi-Square Test Statistic Value	42.64398	0.008	**
Degrees of Freedom	23		
Number of Observations	314		
**** Indicates statistical significance at the 5 percent level or better**			
*** Indicates statistical significance at the 10 percent level or better**			

DISCUSSION AND CONCLUSIONS

The primary goal of this paper was to present a method demonstrating how information entropy measures can be used to evaluate the quantity of information provided in survey responses. In doing so, it is possible to examine some of the fundamental assumptions underlying survey design. The study data came from an established sport related survey, which was administered during a very large amateur basketball event. Nine items covering basic activities or experience that underlie individual values were chosen as the primary items upon which information entropy was identified. The empirical findings were twofold. First, most respondents, when characterized by traditional categories such as age, gender, previous experience attending the sporting event, and location of primary residence, provided responses that contained no statistically different quantities of information. This implies that the survey design (which was not stratified and treated all responses across respondents) was generally appropriate.

Second, the empirical results found that individuals who responded in a particular way to a small number of perception questions did, in fact, provide different quantities of information to the survey items regarding individual values. More specifically, the Tobit regression results suggest that individuals who disagreed with the statement that they get caught up in others' feelings, individuals who agreed or strongly agreed that they can intuit others' feelings (i.e., express empathy) and who disagreed with the contention that the sporting event contributes meaningfully to the local community all provided significantly higher information content (relative to the corresponding omitted category for each of these variables) holding the other specified regressors constant. Clearly, people who view themselves as empathetic, yet are unswayed by others' emotions, provided more informative responses about their own set of values. One obvious explanation for these findings is that those individuals who fall into the aforementioned categories likely have developed a more definitive sense of their own values, and thus are more likely to provide different (and more informative) responses compared to those who have not self-identified a strong sense of personal attitudes and beliefs.

The results also present a policy recommendation to the authors of the survey. If the survey is administered during future Hoopfest tournaments, it may be useful to add one or more pre-screening questions to first identify individuals who are likely to provide different quantities of information in their responses. Such pre-screening may allow the authors to exploit information differences and make better use of the extra information contained in these responses. This may also allow the researchers to further explore the attitudes and beliefs of these respondents and more fully uncover the mechanisms through which they inform these individuals' more informative responses about their values.

While the current study serves as an interesting examination of information entropy assessment, it is by no means exhaustive, and future work is necessary to explore how information entropy may impact survey design and administration. The survey in this particular study was administered to respondents who, by and large, gave equally informative responses. Hence, we can say little about how information entropy might shape survey design and administration when different groups of respondents provide substantial and meaningfully different quantities of information in their responses, e.g. empathetic versus non-empathetic respondents in prosocial behavior research. It may also be the case that if a fundamentally different survey was administered to the same population using the same administration techniques, the results of the analysis might change, e.g. respondents are allowed to discriminate among reasons for the event's perceived importance. Future applications of this methodology to different surveys and populations would therefore provide greater insight into the usefulness of information entropy assessment to marketing researchers.

Another limitation of the information entropy construct is that it simply identifies potential opportunities to revise a survey to more effectively utilize responses. Those opportunities may arise through biases, such as leniency errors or faming biases. The use of information entropy will not identify which biases are causing the information loss. Rather, the researcher must interpret entropy results within the context of proper survey design techniques to identify these biases.

A third limitation of the analysis is that our information entropy measure is more difficult to apply in survey-based research designs where responses are not mutually exclusive or collectively exhaustive. Thus, questions that require respondents to rank order alternatives, or that allow respondents to avoid answering a question (i.e., "not applicable" or "do not know" responses) may not provide information that (without further transformations) facilitates the construction of an entropy measure. Future research is necessary to develop information entropy measures that fully characterize the quantity of information contained in these alternative question designs.

Lastly, it is important to note that the use of information entropy is *not* intended to serve as an exhaustive measure of a survey's utility. Rather, it serves as an *additional* tool that supplements existing measures of reliability and validity currently in use by marketing researchers. Using the information entropy along with other metrics should empower researchers to better triangulate the overall effectiveness of survey-based research, including both design and implementation features of that research. Future research is necessary to determine the appropriate set of measures best suited for such triangulation analyses.

ENDNOTES

1. If the response scale is continuous (for example, if survey items ask for written responses based on time, money, etc.) it may be convenient to enforce discretization of the continuous variable into a fixed set of categories which cover the range of response. Of course, such recoding will likely reduce the quantity of information available within the variable.
2. These last two assumptions may be relaxed in certain circumstances, although doing so complicates the interpretation of any resulting information entropy calculations.
3. Normally, $L_i = L$ for every $i = 1,\ldots,n$; that is, each respondent answers the same number of survey items. While we assume this holds in our analysis, this need not be the case.

REFERENCES

Bas, O. & Grabe, M. E. (2015). Emotion-provoking personalization of news: Informing citizens and closing the knowledge gap? *Communication Research*, *42*(2), 159-185.

Bearden, W.O. & Netemeyer, R.G. (1999). *Handbook of marketing scales: Multi-item measures for marketing and consumer behavior research.* Thousand Oaks, CA: Sage Publications.

Berger, A. (2013). *Media and communication research methods: An introduction to qualitative and quantitative approaches*. Thousand Oaks, CA: Sage Publications.

Bozman, C. S., Kurpis, L.V. & Frye, C. (2010). Hoopfest: Using longitudinal economic impact data to assess the success of a strategic reorientation. *Sport Management Review*, *13*, 65-81.

Cottrell, R. R. & McKenzie, J. F. (2005). *Health promotion education research & methods: Using the five-chapter thesis/dissertation model.* Sudsbury, MA: Jones and Bartlett.

Cox, E. P. (1980). The optimal number of response alternatives for a scale: A review. *Journal of Marketing Research, 17*(4), 407-422.

Dahl, F. & Osteras, N. (2010). Quantifying information content in survey data by entropy. *Entropy, 12*(2), 161-163.

Dillman, D. A. (2000). *Mail and internet surveys: The tailored design method.* New York, NY: John Wiley and Sons.

Dowling, T. C. (2015). Randomized control trials. Chapter 4 in *Principles of Research Design and Drug Literature Evaluation*, R. R. Aparasu and J. P. Bentley, eds. Burlington, MA: Jones and Bartlett.

Eisend, M. (2015). Have we progressed marketing knowledge? A meta-meta-analysis of effect sizes in marketing research. *Journal of Marketing*, *79*(3), 23-40.

Fiedler, B. & Bebber, R. (2013). An international regulatory clinical trial comparative. *International Journal of Pharmaceutical and Healthcare Marketing, 7*(2), 199-212.

Fiedler, B., Bebber, R. & Oetjen, R. (2013). A regulatory clinical trial primer: Distribution challenges and policy implications of modern vaccine development. *International Journal of Pharmaceutical and Healthcare Marketing, 7*(1), 45-57.

Finn, S. and Roberts, D.F. (1984). Source, destination, and entropy: Reassessing the role of information theory in communication research, *Communication Research, 11*(4), 453-476.

Friesner, D., Khayum, M. & Schibik, T. (2013). Characteristics of the information content in business sentiment surveys. *American Journal of Business, 28*(1), 19-37.

Golan, A. (2006). Information and entropy econometrics – A review and synthesis. *Foundations and Trends in Econometrics, 2*(1-2), 1-145.

Golan, A., Judge, G. & Miller, D. (1996). *Maximum entropy econometrics: Robust estimation with limited data.* New York, NY: John Wiley and Sons.

Greene, W. (2000). *Econometric analysis, 4th Edition*. Upper Saddle River, NJ: Prentice Hall.

Hoffman, M. L. (2000). *Empathy and moral development* (1st ed.). Cambridge, UK: Cambridge University Press.

Homer, P.M. & Kahle, L.R. (1988). A structural equation analysis of the value-attitude-behavior hierarchy. *Journal of Personality and Social Psychology*, *54*(3), 638-646.

Hsia, H. J. (1988). *Mass communications research methods: A step-by-step approach*. Hillsdale, NJ: Lawrence Erlbaum.

Jaynes, E. (1957). Information theory and statistical mechanics. *Physics Review*, *106*(4), 620-630.

Jaynes, E. (1982). On the rationale of maximum-entropy methods. *Proceedings of the IEEE*, *70*(9), 939-952.

Kahle, L. R. (1983). *Social values and social change: Adaptation to life in America*. New York, NY: Praeger.

Kurpis, L. V., Bozman, C. S. & Kahle, L. R. (2010). Distinguishing between amateur sport participants and spectators: The list of values approach. *International Journal of Sport Management and Marketing, 7*(3/4), 190-201.

Lijiang, S. (2010). Mitigating psychological reactance: The role of message-induced empathy in persuasion. *Human Communication Research*, *36*(3), 397-422.

Mackert, M., Guadagno, M., Mabry, A. & Chilek, L. (2013). DTC drug advertising ethics: laboratory for medical marketing. *International Journal of Pharmaceutical and Healthcare Marketing, 7*(4), 374-390.

Masri, M., Ramirez, B., Popescu, C. & Reggie, E. (2012). Contract research organizations: An industry analysis. *International Journal of Pharmaceutical and Healthcare Marketing, 6*(4), 336-350.

Massi Lindsey, L. L., Kimo Ah, Y. & Hill, J. B. (2007). Anticipated guilt as motivation to help unknown others. *Communication Research*, 34(4), 468-480.

Merolla, A. J., Zhang, S. & Sun, S. (2013). Forgiveness in the United States and China: Antecedents, consequences, and communication style comparisons. *Communication Research*, *40*(5), 595-622.

Rucinski, D. (2004). Community boundedness, personal relevance, and the knowledge gap. *Communication Research*, *31*(4), 472-495.

Schibik, T., Khayum, M. & Friesner, D. (2012). The quantity of unique information in student ratings of instruction. *Atlantic Economic Journal, 40*(2), 221-223.

Schnell, L. (2014). Hoopfest, the world's largest 3-3 tourney, turns 25 this weekend. Retrieved from http://www.si.com/college-basketball/2014/06/27/

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 37-423.
Slack, M. K. & Baidoo, B. (2015). Bivariate analysis and comparing groups. Chapter 10 in *Principles of Research Design and Drug Literature Evaluation*, R. R. Aparasu and J. P. Bentley, eds. Burlington, MA: Jones and Bartlett.
Smith, S. M. & Albaum, G. S. (2005). *Fundamentals of marketing research.* Thousand Oaks, CA: Sage Publications.