Leisure Sciences

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ulsc20

When Do Normative Standards Matter Most? Understanding the Role of Norm Strength at Multiple National Park Settings

Walter F. Kuentzel a, Daniel Laven b, Robert E. Manning a & William A. Valliere a

a Rubenstein School of Environment and Natural Resources, University of Vermont, Burlington, VT, USA
b Marsh-Billings-Rockefeller National Historic Park, National Park Service,

Available online: 07 Mar 2008

To cite this article: Walter F. Kuentzel, Daniel Laven, Robert E. Manning & William A. Valliere (2008): When Do Normative Standards Matter Most? Understanding the Role of Norm Strength at Multiple National Park Settings, Leisure Sciences, 30:2, 127-142

To link to this article: http://dx.doi.org/10.1080/01490400701881374

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
When Do Normative Standards Matter Most? Understanding the Role of Norm Strength at Multiple National Park Settings

WALTER F. KUENTZEL
Rubenstein School of Environment and Natural Resources
University of Vermont
Burlington, VT, USA

DANIEL LA VEN
Marsh-Billings-Rockefeller National Historic Park
National Park Service

ROBERT E. MANNING
WILLIAM A. VALLIERE
Rubenstein School of Environment and Natural Resources
University of Vermont
Burlington, VT, USA

Research has demonstrated variation in the prevalence, importance, and stability of normative standards across different settings and activities. None of these studies, however, has directly used the concept of norm strength to help explain this variation. This study used norm strength to explore variation in normative standards at 52 locations in 13 U.S. national parks. The analysis measured and supported five dimensions of norm strength: intensity, consensus, certainty, preference consistency, and management consistency. No significant correlations were found between the five indicators indicating conceptually distinct dimensions of the norm strength concept. The findings support the notion that normative standards matter more at some places and the norm strength concept can help managers better understand these differences.

Keywords discriminant analysis, encounter norms, national parks, norm strength, social carrying capacity

The normative standards framework has been widely applied to management concern about recreational crowding over the past 30 years (Manning, 2007). Crowding was originally conceived as a relationship between use levels and visitor satisfaction. Since the 1970s, however, crowding has been viewed as a perception or social evaluation about how many...
people are “too many.” Management frameworks such as Limits of Acceptable Change (Stankey et al., 1985), Visitor Impact Management (Kuss, Graefe, & Vaske, 1990), and Visitor Experience and Resource Protection (Manning, Lime, & Hof, 1996) each rely on collective judgments by recreational visitors about how many people are too many in a given recreational setting and context. These frameworks use the concept of normative standards, or evaluations about appropriate or acceptable conditions that are generally shared by a population of recreational visitors.

Like any prominent framework, the normative standards concept has come under scrutiny over the years. The norm prevalence debate (Donnelly, Vaske, Whittaker, & Shelby, 2000; Roggenbuck, Williams, Bange, & Dean, 1991) considered the extent to which normative standards exist in recreational settings by analyzing the proportion of a sample that was able to specify an encounter norm. The issue of existing conditions and their effect on crowding and normative standards (Laven, Manning, & Krymkowski, 2005; Stewart & Cole, 2001) disputed the extent to which normative standards are relevant by examining the influence of onsite cues on the way people formulate encounter norms. The issue of change versus stability (Bacon, Manning, Johnson, & Vande Kamp, 2001; Kuentzel & Heberlein, 2003) questioned encounter norms stability over time given the rapidly changing social and managerial conditions facing recreation resource managers.

The common thread that binds these issues is that they are fundamentally questions about norm strength. Norm strength is a characteristic of a social norm that reflects the degree of collective agreement about an evaluative standard. For example, many people would probably agree that jet skis are inappropriate on water bodies in wilderness areas, and many would work to ban their presence and report subsequent violations. Fewer people might agree that hunting in a wilderness area is inappropriate, and visitor response to encounters with hunters could vary widely. The norm strength concept acknowledges that variation exists in the prevalence, importance, and stability of normative standards at different settings and social contexts.

Jackson (1965), whose return potential curve first provided a way to measure collective normative standards, suggested two dimensions of norm strength: intensity and crystallization. Intensity is a measure of collective importance, while crystallization is a measure of collective consensus. Other dimensions of norm strength, however, could account for variation in normative standards. Norm prevalence may be the most fundamental dimension of norm strength. The certainty about a normative standard, or norm congruence or consistency between preferred conditions and acceptable conditions could also have an influence.

We use the concept of norm strength to examine factors that might make normative standards at some settings more prevalent, more consequential, and more stable than at other settings. The analysis expands the norm strength concept by exploring additional dimensions besides the traditional intensity and crystallization measures. The analysis is based on data from 52 locations in 13 U.S. national parks used to investigate variation in the strength of normative standards across different settings and activities as well as different measurement methods.

Conceptual Challenges

Conceptual challenges to the normative standards framework have drawn from data sources that offer mixed evidence. One of the earliest challenges questioned the relevance of encounter norms. Roggenbuck et al. (1991) sampled a group of whitewater rafters on a high-use stretch of the New River in West Virginia. They asked respondents to report the greatest acceptable number of river encounters on three types of whitewater experiences: wilderness/backcountry whitewater, scenic/frontcountry whitewater, and social/high-use
whitewater. In each setting, less than half and as few as 29% of the sample were able to specify an encounter norm. Shelby and Vaske (1991), however, found that at least three-fourths of the people on several Oregon rivers were able to specify an encounter norm. In a later study using 30 studies and 56 evaluation contexts, Donnelly et al. (2000) found that norm prevalence, or the ability to specify a norm, ranged from a low of 29% from the New River rafters study to a high of 97% among Colorado bow elk hunters. Their study used multiple data sets to demonstrate that normative standards matter, but there can be variation in norm prevalence depending on the setting (i.e., frontcountry/backcountry), type of encounter (i.e., conflict/no conflict), and questionnaire response format (i.e., 2-category/3-category).

A second challenge to the normative standards framework questioned the relevance of encounter norms because of the seemingly arbitrary way people develop standards. The existing conditions critique argues that when people first visit a recreational site, they have little knowledge and vague expectations about appropriate use levels. As a result they use existing social conditions as a cue for judging the normative standard for the site, and may evaluate whatever condition they encounter as acceptable (Roggenbuck et al., 1991; Stewart & Cole, 2001). Repeat visitors to a recreational setting may also use site cues to adjust their normative standard. Similarly, the product shift literature (Kuentzel & Heberlein, 1992; Shindler & Shelby, 1995) argues that people often respond to unexpected conditions by changing their standards to correspond more closely with their experiences. Laven et al. (2005), however, tested the strength and variability of the relationship between normative standards and existing conditions at 37 locations within 11 U.S. national parks. A variety of normative standards were measured including encounters, campsite impacts, trail impacts, and litter/graffiti. Their results found almost no relationship between normative standards and what visitors evaluated as current or existing conditions at each of the locations. They also found no relationship when the sample was split between first-time and repeat visitors. This study used multiple data sets to show that normative standards in national park settings are derived from something more than just onsite cues.

A third challenge focuses on whether normative standards tend to be stable or whether they are subject to change as social and managerial conditions evolve. Recent longitudinal research has generated seemingly conflicting results. Kuentzel and Heberlein (2003) found that encounter norms changed among recreational boaters at the Apostle Islands National Lakeshore between 1975 and 1997. Visitor use more than doubled between 1975 and 1985. Yet, encounter norms increased in 1985 and visitors in 1985 felt less crowded. Use-levels and encounter norms remained stable between 1985 and 1997, but visitors felt more crowded. In a study of backcountry visitors to Denali National Park, however, Bacon et al. (2001) reported evidence for norm stability over a 22-year interval. Respondent encounter norms, reported encounters, and perceived crowding scores did not differ significantly between 1978 and 2000. Other studies found mixed and inconclusive results about norm stability over time (Shelby, Bregenzer, & Johnson, 1988; Cole, Watson, & Roggenbuck, 1995). While there is not enough longitudinal data in the field to settle the stability versus change issue, it appears that the stability of normative standards may vary depending on settings, activities, and managerial contexts.

Considered together, three lessons may be taken from these conceptual and empirical debates about normative standards in recreation. First, the analysis of multiple data sets has been an effective method for casting a broader and clearer perspective on the important issues raised by proponents and critics of this framework. Second, the results have shown that the normative standards framework is a fairly robust approach to understanding visitor expectations in spite of obvious variation in how broadly those standards may be shared across a user population (Heywood, Manning, & Vaske, 2002). Third, the variation in norm prevalence, importance, and stability demonstrated by these debates are fundamentally
questions of norm strength. This variation calls for a better understanding of the norm strength concept, and an analysis of the conditions under which normative standards are more or less strongly held.

**Norm Strength**

The concept of norm strength characterizes the overall degree of collective agreement and importance of a normative standard. Some norms matter more than others, and some norms exert a stronger influence on behavior than others. A norm against hiking with unleashed dogs on the Appalachian Trail is probably not as clearly defined or accepted as is a norm against shooting unleashed dogs on the Appalachian Trail. Norm strength is also related to behavioral outcomes. When violated, strongly held normative standards can elicit behavioral reactions such as displacement or sanctions (Heywood, 2002) and onsite conflict (Donnelly et al., 2000). Most recreation research, however, has focused on measuring standards and paid somewhat less attention to measuring norm strength.

Recreation researchers have followed Jackson’s (1965) original *return potential curve* formulation to measure normative standards and norm strength. Jackson’s norm curve is derived from an X-axis, which includes various levels of some condition (e.g., number of people encountered, number of unleashed dogs encountered), and a Y-axis, which represents an evaluative scale from positive at the top to neutral in the middle to negative at the bottom. Respondents evaluate the acceptability of each level on the X-axis. The mean acceptability scores for each level are then plotted on the graph. A line is drawn from each mean score to plot the curve. In recreation research, the point on the X-axis at which encounter evaluations fall below the neutral line is sometimes used as a guide to establish visitor carrying capacities. The structural properties of this measurement approach allow researchers to characterize an optimal condition (i.e., the highest acceptable point of the curve), a range of acceptable conditions (i.e., points along the curve that lie above the neutral line), the importance of the standard, and the overall agreement about the standard (Donnelly et al., 2000; Manning, 1999; Shelby & Heberlein, 1986).

These last two structural properties of the norm curve were conceptualized as dimensions of norm strength. Intensity measures the importance of a normative standard and is measured by the distance from the most positive and the most negative evaluations on the curve. More intensely held norms display “steeper” lines, while “flatter” lines reflect less intensively held norms, or more adaptable or accommodating evaluation standards. Crystallization measures the consensus or strength of agreement of the normative standard and is sometimes operationalized as the average of all standard deviations for each evaluation. Smaller average standard deviations are indicators of greater consensus, while greater average standard deviations reflect less consensus and more ambivalence.

Most recreation research has not strayed from Jackson’s (1965) two dimensions of norm strength. Most studies treat these two dimensions as descriptive components of norm structure (Manning, 1999). Few studies, however, have specifically targeted the norm strength concept as the focus of analysis. One exception is Heywood’s (2002) research on urban park littering and trail etiquette. Heywood incorporated the role of sanctions into his normative framework and argued that norm strength was best characterized as the intersection between a cognitive component of obligation to a standard, and an emotional component of embarrassment or guilt when that standard is not followed. When obligation is high, norm consensus is also high. When embarrassment and guilt are high, norm intensity is also high. Heywood’s data demonstrate how the cognitive component (obligation/consensus) and the emotional component (intensity) can be cross-tabulated to measure norm strength and norm prevalence. A second exception is Donnelly et al.’s. (2000) study of norm prevalence.
When Do Normative Standards Matter Most?

their study, norm prevalence is one more component of norm structure along with the optimum condition, the range of acceptable conditions, consensus, and intensity. Following this logic, norm prevalence could be considered another dimension of norm strength. From these studies, it is easy to see how there might be more dimensions to norm strength than just consensus and intensity.

Beyond Consensus and Intensity

Our study incorporated three additional dimensions of norm strength suggested by the structural properties of the norm curve. First, norm certainty reflected the degree to which a sample agrees that a specific standard is inappropriate. For example, a satisfying tennis match can be played with two or four people. However, a fifth player on the court turns the game into something different. This dimension is conceptually the opposite of norm prevalence. Prevalence measures the ability of a group to articulate a standard, while certainty measures the ability of a group to articulate a specific threshold. Norm curves with high certainty will display a steeply falling line between one positively evaluated level and the next negatively evaluated increment.

Next, the analysis constructed two measures of norm consistency. Consistency is the degree to which the sample’s preferences match their normative standards. The first dimension was preference consistency, which reflects the correspondence between the condition people say they prefer and their evaluation of that condition on an acceptability scale. Similarly, management consistency reflects the correspondence between the condition people say requires some management intervention and their evaluation of that condition on an acceptability scale. Preference consistency was measured by calculating the distance between the sample’s preferred condition and the point on the norm curve that falls below the neutral line into the negative evaluation range. Management consistency was measured by calculating the distance between the average condition that requires management intervention and the point on the norm curve that falls below the neutral line.

In sum, this study created five dimensions of norm strength: intensity, consensus, certainty, preference consistency, and management consistency. In exploring multiple dimensions of any construct, questions that immediately arose were whether or not the dimensions are distinct, or are they simply measuring a single or smaller group of constructs? Consequently, we analyzed the correlation matrix for the five norm strength indicators. We hypothesized that each of the five measures was uncorrelated and conceptually distinct dimensions of the norm strength concept (H1).

Norm Strength Variation

Understanding variation in norm strength requires the analysis of multiple data sets. In cross-sectional studies, norm strength measures provide descriptive statistics to assist managers in evaluating the importance and validity of a normative standard. The greater value of norm strength measures is their comparative ability across different studies, and their ability to help managers and researchers understand the conditions under which the overall importance of a normative standard might vary. The literature suggests at least four conditions. Donnelly et al. (2000) used multiple data sets to compare norm prevalence between different settings, activities, types of encounter, and types of response format. The findings about norm change and stability also suggest a setting effect. Bacon et al.’s. (2001) study found norm stability at a backcountry setting, while Kuentzel and Heberlein’s (2003) study found norm change at a frontcountry setting where marinas and docks forced visitors together more frequently. Together these studies indicated that features of the setting such as location, resources, and activities are related to variation in norm strength.
Our study compared the five dimensions of norm strength with different setting characteristics. Past research has only differentiated two setting characteristics: backcountry and frontcountry settings. In our study, however, we followed Andereck and Becker (1993) in differentiating between backcountry, frontcountry, and urban proximate settings among the 52 national park settings. In the U.S. National Park Service, setting characteristics serve as de facto activity characteristics. Backcountry settings include dispersed activities such as hiking or camping, frontcountry settings are the primary attractions where the National Park Service directs most visitors, and urban proximate settings are places close to urban population centers or are the built features of the park such as museums or visitor centers. We hypothesized that norm strength will be different at each of the three types of settings (H2). Similar to Donnelly et al. (2000), we also hypothesized that normative standards will be strongest at backcountry settings and weakest at urban proximate settings (H3).

Lastly, the analysis compared norm strength among studies that used different measurement techniques. Donnelly et al. (2000) found significant differences in norm prevalence depending on the response format that was used in the questionnaire. In addition, Manning, Valliere, Wang, and Jacobi (1999) measured normative standards at Acadia National Park using both visual measurement approaches and numeric measurement approaches, and found greater norm consensus with visual measurement. In this study, we extended this analysis to 12 other national parks and compared norm strength between studies that used visual methods and studies that used numeric methods. Similar to Manning et al., we hypothesized that norm strength will be stronger among those studies that used visual/photographic methods to measure normative standards (H4).

Methods

Study Locations

This study used data from 52 different locations in 13 U.S. national parks ranging from visitor centers to primary attractions to backcountry settings. The data were collected by researchers at the Park Studies Laboratory at the University of Vermont over an 8-year period from 1995 to 2003 (see Table 1). These data came from studies that were part of a long-term research agenda on normative standards. Although the study locations were not representative of all U.S. national parks, they offered a diversity of settings for exploring variation in norm strength.

The analysis categorized each park setting into one of three categories: backcountry settings, frontcountry settings, and urban proximate or built settings. The backcountry settings included regions of the parks that typically require some form of nonmotorized travel (i.e., mostly hiking) and often an overnight camping stay. These locations included Isle au Haut in Acadia National Park, and the backcountry trail systems of Grand Canyon National Park and Zion National Park. The frontcountry settings included the park’s main attractions or where park information services direct most visitors (e.g., the carriage roads in Acadia National Park, Delicate Arch in Arches National Park, Thurston Lava Tube in Hawaii Volcanoes National Park, or Bridalveil Fall in Yosemite National Park). Finally, the urban proximate or built settings included either parks located close to an urban center (e.g., Boston Harbor Islands National Recreation Area) or built settings through which many park visitors pass such as visitor centers (e.g., Haleakala National Park), museums (e.g., Mesa Verde National Park), or historic structures (e.g., Alcatraz Island Prison Cellhouse).

Measurement

Norm curves. A primary goal of each study in Table 1 was to measure encounter norms at locations where managers suspected a crowding problem. Across the different studies,
TABLE 1  Frontcountry, Backcountry, and Urban Proximate Study Locations

<table>
<thead>
<tr>
<th>Front Country</th>
<th>Back Country</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acadia National Park</strong></td>
<td><strong>Acadia National Park</strong></td>
</tr>
<tr>
<td>● Carriage roads</td>
<td>● Isle-au-Haut visitors</td>
</tr>
<tr>
<td>● Scoodic Peninsula (Tour road)</td>
<td>● Isle-au-Haut residents</td>
</tr>
<tr>
<td>● Scoodic Peninsula (Scoodic Point)</td>
<td>● Threshold and Primitive trails</td>
</tr>
<tr>
<td>● Scoodic Peninsula (Fraser Point)</td>
<td>● Corridor trails</td>
</tr>
<tr>
<td><strong>Arches National Park</strong></td>
<td>● Rim trails</td>
</tr>
<tr>
<td>● Delicate Arch</td>
<td>● Virgin River Narrows – onsite</td>
</tr>
<tr>
<td>● Windows</td>
<td>● Virgin River Narrows – Mail back/no photos</td>
</tr>
<tr>
<td>● Devil’s Garden</td>
<td>● Virgin River Narrows – Mail back/photos</td>
</tr>
<tr>
<td><strong>Blue Ridge Parkway</strong></td>
<td>● Grotto</td>
</tr>
<tr>
<td>● Vehicles on road</td>
<td>● Weeping Rock</td>
</tr>
<tr>
<td>● Overlooks</td>
<td>● Canyoneering Backcountry</td>
</tr>
<tr>
<td><strong>Haleakala National Park</strong></td>
<td>(continued on next page).</td>
</tr>
<tr>
<td>● Red Hill Overlook – outside</td>
<td></td>
</tr>
<tr>
<td>● Red Hill Overlook – inside</td>
<td></td>
</tr>
<tr>
<td><strong>Hawaii Volcanoes National Park</strong></td>
<td></td>
</tr>
<tr>
<td>● Thurston Lava Tube trail</td>
<td></td>
</tr>
<tr>
<td>● Thurston Lava Tube</td>
<td></td>
</tr>
<tr>
<td>● Chain of Craters parking lot</td>
<td></td>
</tr>
<tr>
<td>● Chain of Craters Lava Flow</td>
<td></td>
</tr>
<tr>
<td><strong>Kenai Fjords National Park</strong></td>
<td></td>
</tr>
<tr>
<td>● Exit Glacier – Upper Loop trail</td>
<td></td>
</tr>
<tr>
<td><strong>Mesa Verde National Park</strong></td>
<td></td>
</tr>
<tr>
<td>● Balcony House</td>
<td></td>
</tr>
<tr>
<td>● Cliff Palace</td>
<td></td>
</tr>
<tr>
<td>● Long House</td>
<td></td>
</tr>
<tr>
<td>● Sun Point Overlook</td>
<td></td>
</tr>
<tr>
<td>● Spruce Tree House</td>
<td></td>
</tr>
<tr>
<td>● Trail to Spruce Tree House</td>
<td></td>
</tr>
<tr>
<td>● Step House</td>
<td></td>
</tr>
<tr>
<td><strong>Muir Woods National Monument</strong></td>
<td></td>
</tr>
<tr>
<td>● Muir Beach</td>
<td></td>
</tr>
<tr>
<td>● Muir Woods Trails - exit</td>
<td></td>
</tr>
<tr>
<td>● Muir Woods - exit</td>
<td></td>
</tr>
<tr>
<td><strong>Yosemite National Park</strong></td>
<td></td>
</tr>
<tr>
<td>● Trail to Bridalveil Falls</td>
<td></td>
</tr>
<tr>
<td>● Base of Bridalveil Falls</td>
<td></td>
</tr>
<tr>
<td>● Glacier Point</td>
<td></td>
</tr>
<tr>
<td>● Trail to Mirror Lake</td>
<td></td>
</tr>
<tr>
<td>● Trail to Vernal Falls</td>
<td></td>
</tr>
<tr>
<td>● Trail to Yosemite Falls</td>
<td></td>
</tr>
<tr>
<td>● Base of Yosemite Falls</td>
<td></td>
</tr>
</tbody>
</table>

(continued on next page).
sampling procedures varied. Most studies were conducted onsite, but three of the 52 studies used mailed questionnaires. Protocols for measuring norms also varied. Forty-five studies asked people to evaluate encounter conditions represented in photographs, while seven studies asked people to evaluate numeric encounter conditions. Nevertheless, each study used the logic of Jackson’s (1965) return potential curve to measure encounter norms. Respondents were presented with between five to 11 different density levels. They were asked to evaluate the acceptability of each density level on a 9-point scale where -4 = “very unacceptable” and +4 = “very acceptable.” Researchers then plotted the mean acceptability rating for each encounter level. Norm curves were calculated using data from each of the 52 national park locations. Norm strength indicators were derived from each of the 52 norm curves.

Norm strength. We constructed five measures of norm strength. The first measure, intensity, directly followed Jackson’s (1965) formulation. He defined intensity as the numeric distance between the highest acceptability score and the lowest acceptability score, or the difference between the highest and lowest point of the norm curve. Norms are held with greater intensity when certain encounter levels receive uniformly strong favorable evaluations and certain other encounter levels receive uniformly strong negative evaluations. For instance, most wilderness hikers would generally rate 0 encounters as “very acceptable” with a mean of approximately +4. Most hikers would rate 60 people very negatively (either −3 or −4). Few people would deviate from these more extreme responses. To measure consensus, we averaged the standard deviations of each encounter condition evaluated. The smaller the average standard deviation, the greater the consensus about acceptable encounter levels at a given site.

The third dimension of norm strength, certainty, measured a threshold effect. Evidence for a threshold (e.g., the difference in acceptability between 4 and 5 people on a tennis court) will occur in a norm curve that shows a significant change in acceptability between two adjacent units on the X-axis. This analysis measured certainty using a fold-over technique suggested by Schuman and Presser (1981), which averages the distance from the mid-point
When Do Normative Standards Matter Most?

of the acceptability scale for each condition evaluated. The key feature in this measure was the steepness of the norm curve between two adjacent encounter levels. Loglinear shaped curves with strongly positive encounter evaluations up to a certain threshold and strongly negative encounter evaluations beyond that threshold would generate higher certainty scores. Conversely, flatter norm curves with modest positive and negative evaluations across the different encounter levels would generate lower certainty scores.

The fourth indicator of norm strength was preference consistency. Preference consistency was measured by calculating the difference between the sample’s mean preferred encounter condition and the point at which the norm curve crossed below the mid point of the acceptability scale (i.e., the minimum acceptable standard). In most outdoor recreation settings, fewer encounters would be the preferred condition. By extension, when the numeric distance between the preferred condition and the minimum acceptable standard is low, people are unwilling to tolerate encounter levels that deviate from their preferences. Conversely, when the distance between the preferred condition and the minimum acceptable standard is high, people are willing to tolerate encounter conditions that are greater than their preferences. In sum, low values reflect stronger preference consistency, while high values reflect weaker preference consistency.

The final measure of norm strength was management consistency. Like preference consistency, management consistency was measured by calculating the difference between encounter levels that people believe require management intervention and the minimum acceptable standard. Like preference consistency, a lower score was assumed to be an indicator of stronger management consistency. Once encounter conditions become unacceptable to a sample, they would be more likely to call for management intervention (Manning, 2007). Consequently, the mean encounter condition when people believe that management intervention is necessary should be fairly close to the point where encounters on the norm curve fall below the neutral line into the unacceptable range.

Two data manipulations were required to calculate the preference consistency and management consistency indicators of norm strength. First, to determine the point at which the norm curve crossed the mid point of the acceptability scale, linear interpolation was used. The process is shown in the formula below:

\[ I = n^p + \frac{(n^{hn} - n^p)}{lp - hn} \times lp \]

where \( n^p \) was the number of people in the encounter category with the lowest mean positive acceptability rating, \( n^{hn} \) was the number of people in the encounter category with the highest mean negative acceptability rating, \( lp \) was the lowest positive acceptability rating, and \( hn \) was the highest negative acceptability rating. Second, we standardized encounter levels across studies to create consistency scores that were expressed as a proportion of each study’s acceptability level. To calculate this ratio, the average preference and management intervention scores were divided by the mean acceptability score for each of the 52 samples to produce the two consistency measures.

Analysis

The analysis used the norm data from each study as a single data point and created five norm strength indicators for each of the 52 data sets. The analysis used three strategies to test the study hypotheses. First, a correlation matrix was used to test \( H_1 \), to assess whether the various norm strength dimensions were distinct constructs, or whether each was simply measuring the same construct. Second, a discriminant analysis was used to
test H2. Discriminant analysis demonstrates difference in setting characteristics by using the norm strength indicators as independent variables to “predict membership” in each of the three setting categories. One of the outputs of discriminant analysis was a two-dimensional graphic representation of category differences. An additional output was an assessment of how well each park location was classified into the three categories. Finally, one-way analysis of variance was used to test H3 and H4. This analysis compared mean norm strength differences between the setting characteristics (i.e., backcountry, frontcountry, and urban proximate), and compared mean norm strength differences between measurement techniques (i.e., visual and numeric).

Results

Inter-item Correlations

The results from the correlation matrix showed support for H1, as shown in Table 2. The five norm strength variables were generally not correlated with one another. Only 2 of 10 correlations were significant at p < .05. The consensus dimension was positively correlated with the preference consistency dimension (r = .45). When there was greater variation in the way people evaluated encounter conditions (i.e., larger standard deviations across each encounter condition evaluated), there also was greater preference consistency (i.e., greater distance between one’s preferences and the normative standard). In addition, the certainty dimension was negatively correlated with the management consistency dimension (r = −.34). When people’s evaluations of encounter conditions were more certain, they desired more management intervention when encounter conditions exceeded the normative standard. Overall, however, most norm strength dimensions were not correlated indicating that these items were generally independent and were measures of distinct norm strength dimensions.

Discriminant Analysis

The discriminant analysis provided support for H2. Figure 1 plots the discriminant scores for each of the 52 settings, and shows the group centroids for backcountry, frontcountry, and urban proximate settings. Visually, the norm strength variables appeared to have discriminated well with each group occupying roughly distinct spaces in Figure 1. To characterize this empirically, the Box M statistic tests the null hypothesis that the covariance matrix of each group is equal to the population covariance matrix. Since the Box M statistic for this model was 89.3 (p < .001), the null hypothesis was rejected and we concluded a distinct covariance in each group. The second way to characterize model significance is with the χ² statistic calculated for each discriminant function. The χ² for Function 1 was 66.9 (df = 10,
When Do Normative Standards Matter Most?

$\chi^2$ for Function 2 was 14.3 ($df = 4, p = .006$), which explained 15% of the variance in the model. These findings indicated significant norm strength variation between the three groups in their scores for each discriminant function.

Norm strength differences were also examined by comparing the researcher defined classifications (i.e., frontcountry, backcountry, urban proximate) to model predicted classifications. Table 3 shows that model predictions differed from researcher categories on only eight of the 52 locations (15%). Three of the frontcountry locations (i.e., Acadia carriage roads, Arches/Delicate Arch, and the Blue Ridge Parkway/Vehicle evaluations) were predicted by the model to be backcountry locations. For the Acadia carriage roads, the probability of membership in the frontcountry group was 48% while the probability of membership in the backcountry group was 51%, so the data did not clearly discriminate in this one case. The probability of membership in the backcountry group for Arches/Delicate Arch and the Blue Ridge Parkway/Vehicle evaluations were much higher (64% and 73%, respectively). In addition, the model predicted that the Grand Canyon Corridor trails should be categorized as a frontcountry location rather than a backcountry location. Finally, the model predicted that both Red Hill Overlook locations at Haleakala National Park and both locations at the Chain of Craters at Hawaii Volcanoes National Park should be classified as urban proximate/built locations rather than frontcountry locations. Most of our original

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|}
\hline
A Priori Categorization & Back Country & Front Country & Urban Proximate \\
\hline
\textbf{Back Country} & 11 (91.7\%) & 1 (8.3\%) & 0 \\
\textbf{Front Country} & 3 (9.1\%) & 26 (78.8\%) & 4 (12.1\%) \\
\textbf{Urban Proximate} & 0 & 0 & 7 (100\%) \\
\hline
\end{tabular}
\caption{Cross-tabulation of Researcher and Model Categorizations for 52 National Park Locations}
\end{table}
categorizations of these locations, however, agreed with the model predictions. Overall, the discriminant analysis showed substantive differences in norm strength between visitors at different national park settings.

**Analysis of Variance**

Table 4 provided mixed support for H$_3$. The findings showed that the mean intensity score was significantly higher among the backcountry samples ($M = 5.05$) than for the frontcountry ($M = 2.69$) and urban proximate/built samples ($M = 3.10$). The backcountry samples evaluated fewer encounters as more acceptable and higher encounters as less acceptable. The mean consensus score was significantly higher among the urban proximate/built samples ($M = 2.17$) than the frontcountry ($M = 1.93$) and backcountry samples ($M = 1.98$), which indicated larger average standard deviations for each encounter level. This finding suggested less agreement about appropriate encounter levels among at urban proximate visitors. Table 4 shows no significant differences among any of the samples in their mean certainty scores. Finally, mean norm strength scores differed for both of the consistency measures. The urban proximate/built samples had significantly higher mean scores on the preference consistency measure ($M = 0.57$) than the frontcountry ($M = 0.43$) and backcountry samples ($M = 0.39$). The difference between the preferred encounters and the minimal acceptable encounter level for the urban proximate samples was greater, which suggested these visitors were willing to tolerate more encounters than they preferred. Last, the backcountry samples had significantly higher mean scores ($M = 1.46$) on the management consistency dimension than frontcountry ($M = 1.11$) and urban proximate/built samples ($M = 0.99$). The difference between the minimal acceptable encounter level and the encounter condition in which managers should intervene was greater among the backcountry samples. People in the backcountry saw less need for management intervention.

Finally, the results also provided mixed support for H$_4$ (see Table 5). Stronger intensity scores were evident for the numeric measurement method than the visual measurement record. The distance between the most acceptable encounter condition and the least acceptable encounter condition was greater for numeric methods than visual methods. Table 5 also shows that certainty scores and preference consistency scores did not differ between numeric and visual methods. Two comparisons offered support for H$_4$. First, stronger norm consensus was found when using visual measurement as evidenced by smaller average standard deviations across the different locations. Second, there was also greater management consistency when using visual measurement methods.

**TABLE 4** Comparison of Norm Strength Means between Categories of National Park Locations

<table>
<thead>
<tr>
<th></th>
<th>Back Country</th>
<th>Front Country</th>
<th>Urban Proximate</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intensity</td>
<td>5.05$^b$</td>
<td>2.69$^a$</td>
<td>3.10$^a$</td>
<td>13.82$^*$</td>
</tr>
<tr>
<td>Consensus</td>
<td>1.98$^a$</td>
<td>1.93$^a$</td>
<td>2.17$^b$</td>
<td>5.01$^*$</td>
</tr>
<tr>
<td>Certainty</td>
<td>2.80</td>
<td>2.86</td>
<td>2.91</td>
<td>1.79</td>
</tr>
<tr>
<td>Preference consistency</td>
<td>0.39$^a$</td>
<td>0.43$^a$</td>
<td>0.57$^b$</td>
<td>8.56$^*$</td>
</tr>
<tr>
<td>Management consistency</td>
<td>1.46$^b$</td>
<td>1.11$^a$</td>
<td>0.99$^a$</td>
<td>9.18$^*$</td>
</tr>
</tbody>
</table>

$^*$ $p < .05$.

Superscripts denote category differences that are significant at $p < .05$ using a Duncan range test.
TABLE 5 Comparison of Norm Strength Means between Measurement Methods

<table>
<thead>
<tr>
<th></th>
<th>Visual</th>
<th>Numeric</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intensity</td>
<td>3.05</td>
<td>4.77</td>
<td>8.31</td>
</tr>
<tr>
<td>Consensus</td>
<td>1.95</td>
<td>2.16</td>
<td>7.58</td>
</tr>
<tr>
<td>Certainty</td>
<td>2.87</td>
<td>2.80</td>
<td>1.49</td>
</tr>
<tr>
<td>Preference consistency</td>
<td>0.43</td>
<td>0.47</td>
<td>0.85</td>
</tr>
<tr>
<td>Management consistency</td>
<td>1.11</td>
<td>1.59</td>
<td>20.27</td>
</tr>
</tbody>
</table>

* p < .05.

Discussion

Norm Strength Dimensions

One goal of this study was to expand Jackson’s (1965) conceptualization of norm strength, which included an intensity dimension and a crystallization dimension. Results from the correlation analysis showed the five dimensions of norm strength tested in this study were generally distinct dimensions. In any study where untested constructs are developed, one of the first tasks is validity testing. Are the indicators valid measures of distinct dimensions, or are they simply measures of fewer dimensions, or even a single dimension? The lack of many significant correlations between the five norm strength dimensions supported the expanded framework.

Future studies, however, should refine these indicators and test additional dimensions. The mean certainty scores did not differ between the three activity settings and consequently had little influence on the discriminant analysis. Different types of outdoor recreation sites did not appear to have the same type of threshold effect as does an activity such as tennis. This indicator may need to be compared at a more specific activity level than the more general site level comparisons made in our study. In addition, we predicted that backcountry sites would score higher on management consistency than frontcountry or urban proximate sites. Yet, the difference between the minimal acceptable standard and the encounter level at which visitors advocated management intervention was greatest among the backcountry samples. One would expect that in settings with more strongly held norms, visitors would prefer management intervention fairly close to the point at which encounters begin to negatively affect visitor experiences. It could be, however, that backcountry users are more opposed to management intervention simply because backcountry destinations by definition are free of management interventions. In spite of these caveats, the results demonstrated the utility of an expanded norm strength framework for future research.

Norm Strength Differences

The findings supported the notion that normative standards matter more at some places than at others. The ability of the five norm strength variables to differentiate between visitors at backcountry, frontcountry, and urban proximate/built national park settings was quite good. People who hike, camp, climb, or paddle in backcountry settings have fairly well-defined standards about appropriate conditions, but are generally more averse to management intervention that might alter the dispersed type of recreational opportunities. A backcountry user’s standard of solitude probably has little flexibility. Conversely, people who are found at visitor centers, museums, and other built attractions in the national parks have more ambivalence about appropriate conditions, and are willing to tolerate more encounters than...
they prefer at these sites. Apparently in locations with greater variation in the number of visitors, people can be significantly more flexible about their encounters.

The results echoed Donnelly et al.’s (2000) finding that norm prevalence is greater at backcountry settings than frontcountry settings. The ability to articulate an encounter norm should be closely related to the norm strength concept. If there is little flexibility in the encounter norm among backcountry users, then they would be more likely to identify a standard. Conversely, if a greater tolerance for encounters exists at urban proximate settings, people at these locations would have greater difficulty identifying a standard. Encounter tolerance may also account for variation in the use of existing conditions to articulate a normative standard. People at backcountry settings have less ambivalence about normative standards. Therefore, site cues may have little relevance in articulating a normative standard. Conversely, people at frontcountry or urban proximate settings have relatively more ambivalence about normative standards and, therefore, may be more likely to use existing conditions as cues to identify a normative standard.

The results could also help explain variation in the stability of encounter norms. Encounter norms at Denali National Park did not change between 1978 and 2000 (Bacon et al., 2001). Possibly the expectation for solitude, as legislated by the Wilderness Act, is well understood by most visitors to backcountry areas. Conversely, visitors’ understanding of an encounter norm at a Great Lakes boating destination may be less well understood, particularly at a destination where a variety of social conditions have changed over the years (Kuentzel & Heberlein, 2003). Perhaps the experience of encounters at a Great Lakes boating destination may be different from the experience of encounters in a backcountry setting such as Denali. While user density is a common indicator of the quality of the visitor experience, the way density is manifested could vary from context to context. The quality of the visitor experience can be negatively affected by density at some point and in some way. In the backcountry, the most salient effect is probably encounters. However, in the frontcountry, negative effects may be more about waiting time to access services and facilities, ability to take a picture without someone in the way, or the lack of good anchoring spots. In studying indicators that matter more in backcountry than frontcountry, norm strength is generally higher in the former than the latter.

While the discriminating ability of the norm strength dimensions was good, the analysis of variance showed that the differences between the three settings were not always clear. Backcountry visitors almost always differed from urban proximate/built visitors. However, backcountry visitors differed from frontcountry visitors on only two of five of the norm strength dimensions. The mean scores on the intensity dimension and the management consistency dimension were significantly higher among the backcountry samples than the frontcountry samples. The backcountry samples did not differ from the frontcountry samples on the consensus, certainty, and preference consistency dimensions. Once a visitor leaves the visitor center or the parking lot to explore various parts of the national park, there may be less ambiguity about encounter norms regardless of whether one is visiting one of the park’s main attractions or is visiting a more secluded backcountry setting. Maybe norm strength indicators that rely on averaged measures of response distribution (e.g., averaged standard deviations) are less able to distinguish between backcountry and frontcountry settings. Future research should take a closer look at the response distribution of evaluations across the different encounter levels and experiment with alternative measurement strategies. For example, an alternative measure of consensus might be to average the extremes (i.e., the standard deviation of the most acceptable encounter level with the most unacceptable encounter level).

Finally, the results confirmed Manning et al.’s (1999) finding that differences in norm strength were partly the result of a methods effect. Visual methods that use photographic
When Do Normative Standards Matter Most?

representations of encounters levels generally demonstrated stronger normative standards than methods that use numeric encounter levels presented in a textual format. A greater norm intensity, norm consensus, and management consistency was found when people could see the number of others encountered. This finding does not mean that one method is better than the other. There may be less measurement error when using visual methods and therefore greater consensus and consistency. Yet one can still assume that the greater measurement error in numeric methods is normally distributed and therefore still provides an accurate representation of the normative standard. The results show, however, that measurement method is an effect that must be controlled in a comparative analysis.

Conclusion

This study compared norm strength at multiple national park settings. Norm strength is a concept originally suggested by Jackson (1965) and acknowledged by subsequent recreation researchers, but never fully developed within the social carrying capacity literature. This research used the concept to help understand variation in normative crowding standards. The question of norm strength is an important issue for resource managers who have relied on social carrying capacity models to direct recreation policy. Its application relies on the assumption that the social and environmental standards that visitors apply to their recreation experiences are not the same across different settings and different activities. Understanding the conditions under which normative standards are more or less strongly held can help managers adapt more easily to social change. We also view the norm strength construct as a promising way to help understand some of the questions and discrepancies that are now emerging in the maturing field of crowding and social carrying capacity studies. Future research should continue to develop norm strength indicators and test the reliability and validity of the measures. Research should also incorporate time series indicators of norm strength to verify the link between the norm strength concept and stability and change at different recreational settings.

References


