Maine's Experience Evaluating Wind Energy Development

James F. PALMER

1 Introduction

On April 18, 2008, Governor John Baldacci signed *An Act to Implement Recommendations of the Governor's Task Force on Wind Power Development* (the Act). It establishes a favorable State policy encouraging grid-scale wind energy development in appropriate locations. In particular, it designates a large portion of the state for expedited grid-scale wind energy development. While most environmental impacts within the expedited area are evaluated in the same manner as previously, special provisions are made for scenic impacts.

Since passage of the Act, four visual impact studies for grid-scale wind energy projects have conducted surveys of users at a viewpoint within a scenic resource of state or national significance (SRSNS) where the proposed project's wind turbines are expected to have clearly visibility. At a minimum, respondents were asked to rate the existing view, a photograph of the existing view and a photosimulation of the view with the project turbines using a rating scale of lowest to highest scenic value. Additional questions asked about how the proposed project would affect their recreation enjoyment at the viewpoint, and whether they would return to the area to recreate if the project is built. One survey included a photograph and photosimulation from a second viewpoint. An additional web-based survey for one project also included two impacted viewpoints as well as two additional photographs, but of course could not include ratings of the view in the field.

This paper presents the results of an independent analysis of the data from these user surveys conducted for grid-scale wind energy projects being permitted under the Act's Evaluation Criteria.

2 Methods and Materials

A total of seven viewpoints from four wind projects have been evaluated in the user surveys conducted by Market Decisions and Portland Research Group.

2.1 Perception of Scenic Value and Scenic Impact

Most of the data analyzed here were collected by intercepting people engaged in recreation pursuits at or near the viewpoint being evaluated. The major exception is the web-based survey for Highland, which used a random sample of people from a marketing firm's panel of outdoor recreation participants in northern New England. Survey respondents were shown photographic representations of the existing condition and a photosimulation of the

Buhmann, E., Ervin, S. M. & Pietsch, M. (Eds.) (2013): Peer Review Proceedings of Digital Landscape Architecture 2013 at Anhalt University of Applied Sciences. © Herbert Wichmann Verlag, VDE VERLAG GMBH, Berlin/Offenbach. ISBN 978-3-87907-527-0.

visual condition if the proposed project were built. They were instructed to view these representations from a specific distance, so that they would appear in proper perspective. Scenic evaluations were made using a standard 7-point or 10-point rating scales. In all cases, a rating of 1 is given to the lowest scenic value. A measure of scenic impact is obtained by subtracting the scenic value of the view without the project from the scenic value of the view with the project. However, the meaning of the ratings and the raw impact values are difficult to compare since two different rating scales are used. In addition, the impact value provides no intuitive sense of when the impact is Unreasonably Adverse. While *t*-tests show that these perceived scenic impacts are statistically significant, this is not an indication that the impacts reach the threshold of Unreasonably Adverse. [1]

These shortcomings are common to most scenic impact assessments. However, they can be overcome using the procedures described in the remainder of this section. First the reliability of the data that were collected is considered. Then two possible ways of describing change in scenic value are described: percent change has more intuitive appeal, while effect size has become the preferred way to report findings about change in the scientific literature.

2.2 Visual Simulations

All the images used in these surveys are based on photographs captured with high resolution digital cameras. Each photo was taken using a "normal" lens focal length, though some of the images used in the survey were panoramas created by stitching together two or more normal photos.

2.3 Effect size

The current best practice in scientific analysis and reporting is to use effect size as a way to report the strength of the relationship between the means of two variables measured on the same scale (e.g., APA 2010, p. 33). The statistic used in this paper is HEDGES' (1985) *g*, which estimates the effect size based on the difference between means. Effect size is also comparable across all viewpoints because it is not affected by the rating scale.

STAMPS (1997, 2000) describes how to conduct a study investigating scenic impacts, and how a local Development Review Board might use the effect size results to determine whether or not these impacts are acceptable. After investigating thousands of paired landscape scenes, STAMPS (2000, page 162) has adopted the effect size thresholds suggested by COHEN (1988, pp. 24-27): when $d = 0.2$ it is too small to be noticed, $d = 0.5$ is a medium effect size that is "large enough to be visible to the naked eye," and $d = 0.8$ is large enough to be "grossly perceptible."[2] However, COHEN warns that "the terms 'small,' 'medium,' and 'large' are relative, not only to each other, but to the area of behavioral science or even more particularly to the specific content and research method being employed in any given investigation" (COHEN 1988, p. 25). Fortunately we have measures

1 Statistical significance means the results are unlikely to occur by chance, not that the magnitude of the difference is large enough to be important or even noticeable in the everyday world.

2 "Grossly" has a variety of meanings, but I believe the intent here follows the online Oxford English Dictionary's seventh definition: Indelicately, indecently.

for two of the Act's criteria that can be used to appraise these thresholds for how the scenic impact might affect the enjoyment of using a SRSNS or the continued use of SRSNSs.

3 Results

3.1 Reliability

The first thing to consider is whether the respondents' scenic ratings are reliable – that is if another survey is conducted in a similar manner, how confident should we be that the results would be the same. Reliability can range between 0 and 1. NUNNALLY (1978) states that reliability coefficients of 0.70 or 0.80 are normally acceptable for research purposes, but that reliability should be 0.90 or higher in situations where the measurements are the basis of important decisions.

The intraclass correlation coefficient is the standard measure of reliability for individual respondents, as well as for the group of respondents (PALMER & HOFFMAN, 2001). The individual reliabilities are quite low – they average 0.30 for the intercept surveys and 0.01 for the web survey. However, reliability can be improved by averaging the responses for a group of respondents; normally the more respondents, the higher the reliability. The group reliability for the intercept surveys is quite high (0.91), though the average for the web surveys is still rather low (0.44). In general, the ratings with and without the proposed project from each viewpoint are very reliable for the intercept surveys, though it may be necessary to survey more than 35 respondents to reach reliabilities of 0.90 or higher.

The reliability of the web survey is very low, even though they included 104 respondents who indicated that they recreated in western Maine. In other contexts, web surveys have been shown by others to be an effective tool for public use to evaluate scenic impacts (ROTH 2006, WHERRETT 1999). However, web surveys do not provide an opportunity to experience the context within which the scenic change is viewed (i.e., realism validity). Nor can a web survey provide data useful to estimate the nature, extent and duration of use at a scenic resource, as does a survey that intercepts users at potential viewpoints.

3.2 Scenic Impact

The average effect size from Table 1 for the intercept surveys is -1.21; for the web survey it is -0.27. The result suggests that respondents at the viewpoint are much more sensitive to the potential scenic impact than are the respondents to the web survey.

3.3 Effect on Enjoyment

The Act requires that permitting agencies consider "the potential effect of the generating facilities on the public's… enjoyment of the scenic resource of state or national significance." On a 10-point scale, a mean response of 5.5 indicates that the presence of the turbines would have no effect on the enjoyment of the scenic resource. The mean value for all four studies is 5.055, indicating that on average the presence of the wind turbines would have a slight negative effect. This effect is more negative (and statistically more significant) for the respondents to the Bull Hill and Highland surveys than it is for those to the Saddleback Ridge and Spruce Mountain surveys.

Table 1: Effect Size of Scenic Impact

Case: Project	Pre– $\bar{x}$	Post– $\bar{x}$	Pooled *s*	Effect size
1. Bull Hill (Black Mtn)	8.861	6.003	2.048	-1.395
2. Bull Hill (Donnell Pd)	7.750	6.426	1.853	-0.715
3. Highland (Hikers@9D)	7.514	5.405	1.857	-1.136
4. Highland (Web@9D)	7.971	7.605	1.751	-0.209
5. Highland (Web@4)	7.490	7.308	1.888	-0.097
6. Saddleback Ridge	7.682	5.841	2.606	-0.706
7. Spruce Mountain	6.300	4.950	1.577	-0.856

Using these data, it is possible to demonstrate how one might establish a scenic impact threshold based on when the effect on enjoyment is unacceptable. While only four data points is admittedly quite limited for this task, it is nonetheless still instructive to demonstrate the procedure and consider the results.

The statistical procedure used is linear regression analysis. Fig. 1 shows a scatter plot where the x-axis is effect on enjoyment (or continued use) and the y-axis is effect size. Regression analysis determines the line that best fits the data points in these plots. There is an equation that describes this line in the form of:

$$Y = (b*X) + a \tag{1}$$

"Y" is the dependent variable and the equation will be used to determine unknown values for it. The dependent value effect size. "X" is the independent variable that is used to determine the unknown value of Y for a known value of X. The independent variable is effect on enjoyment (or continued use). "b" is the beta coefficient or slope, and it describes the amount of change in Y for one unit of change in X. "a" is the constant or y-intercept, and it is the value where the line intersects the y-axis when the value of x is zero.

The regression equation that describes the line plotted for effect on enjoyment in Figure 1 is:

$$\text{Effect Size} = (0.7 * \text{Effect on Enjoyment}) - 4.6 \tag{2}$$

There is a strong relationship between effect on enjoyment and effect size (adjusted R^2 = 0.875). Though there are only six data points, such a high relationship strongly suggests that there is a valid connection between scenic impact, as measured effect size, and the effect on enjoyment of a SRSNS.

This equation can be used to determine the values of the effect size for specific values of effect on enjoyment, as shown in Table 2. The Act does not provide guidance for when a negative effect on enjoyment is unacceptable. Must the mean rating be 1.5, or could it be 2.0 or 3.0? However, if 3.0 is established by the permitting agency as the appropriate threshold on a 10-point scale, then the threshold for an Unreasonably Adverse scenic impact would be a size effect of -2.5.

3.4 Continued Use

The Act requires that permitting agencies consider "the potential effect of the generating facilities on the public's continued use… of the scenic resource of state or national significance." On a 10-point scale, a mean response of 5.5 indicates that the presence of the turbines would have no effect on the likelihood that a respondent would continue to use the scenic resource. The mean value for the intercept surveys is 5.504, indicating that the presence of the wind turbines would have essentially no effect on respondents' continued use. However, this effect is substantially greater (and statistically significant) for the respondents to the Highland web survey, who indicate that they would be more likely to return if the project were built ($\bar{x}$ = 7.1).

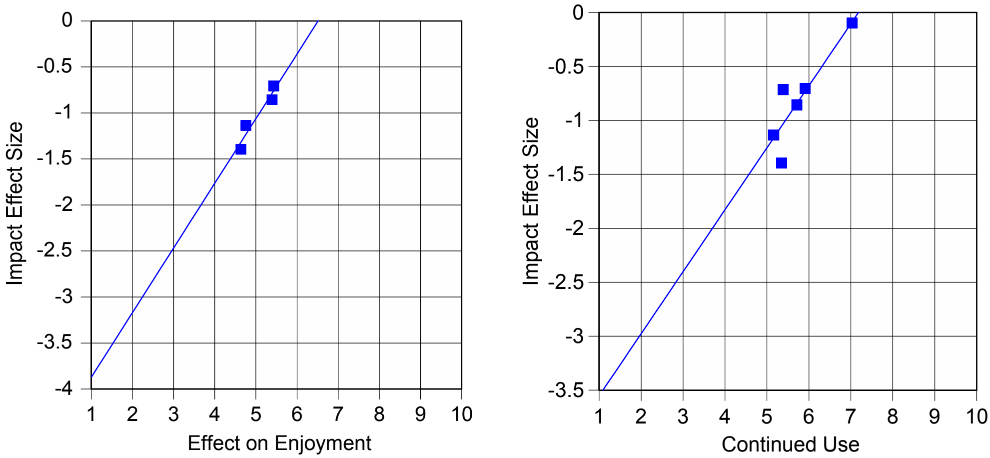

Fig. 1: The relationship between continued use and two measures of scenic impact

Regression analysis is used again to identify the thresholds where the effect size of the scenic impact become Unreasonably Adverse as indicated by users' ratings of whether they would continue to use the SRSNS. The scatter plot and regression line is shown in Fig. 1 There are now severn points and they are more dispersed along the line. However, this analysis must still be treated more as a demonstration of the approach rather than a definitive establishment of thresholds. The regression equation that describes the line plotted in Fig. 1 is.

$$\text{Effect Size} = (0.52 * \text{Continued Use}) - 3.84 \qquad (3)$$

There is a strong relationship between continued use and effect size (adjusted $R^2 = 0.782$). Though there are only seven data points, such a high relationship strongly suggests that there is a valid connection between scenic impact, as measured by effect size, and the continued use of a SRSNS.

This equation can be used to determine the value of the effect size for specific values of continued use, as shown in Table 2. If 3.0 is established as the appropriate threshold on a 10-point scale where the Act considers the negative effect on the likelihood of returning to be Unreasonably Adverse, then the threshold for an Unreasonably Adverse scenic impact would be a -64.4 percent change in scenic value or a size effect of -2.3.

Table 2: Effect Size Values for Thresholds of Effect on Enjoyment and Continued Use

	Extremely Negative (1.5)	**Very Negative (2.0)**	**Negative (3.0)**	**No Effect (5.5)**	**Positive (8.0)**
Effect on Enjoyment	-3.5	-3.2	-2.5	-0.7	1.0
Continued Use	-3.1	-2.8	-2.3	-1.0	0.3

4 Discussion

These results present an opportunity to discuss when the scenic impacts from grid-scale wind energy projects create an Unreasonable Adverse scenic impact. The threshold where the scenic impact becomes Unreasonably appears to be approximately a -2.5 effect size.[3]

This value is admittedly higher than one might expect as the beginning point for such a discussion. For instance, STAMPS (2000) suggests that an effect size threshold of 1.1 be used to identify very large scenic impacts, and the effect size recommended by this study is much larger than that. However, they are based on the judgments of people actually using the affected SRSNSs, and the data appear to be both statistically reliable and valid.

On the other hand, it is acknowledged that there are relatively few data points and these thresholds will need to be recalculated as more surveys are conducted. In addition, it is necessary to include scenic evaluations where the impacts are clearly Unreasonably Adverse. The thresholds suggested here are based on data from viewpoints where users of SRSNS did not think the scenic impact was Unreasonably Adverse.

As more intercept studies targeted to the Act's scenic impact evaluation criteria are conducted a conscious attempt needs to be made to investigate a wider range of users of SRSNSs. The Act requires that the "nature" of the use be considered – hikers may be more or less sensitive to scenic impacts than snowmobilers or people fishing. It also requires that the "duration" of use, and therefore the length of time they are potentially exposed, be considered. To date, all the intercept surveys have evaluated the scenic change at specific viewpoints. It may be that the cumulative exposure to multiple views of wind energy development during a day's outing will result in Unreasonably Adverse threshold levels that are higher or lower than those identified here.

A word of caution about the use of intercept studies may be prudent at this point. While there can be little doubt about the validity of a well conducted intercept study, there is the potential to introduce bias into this method. It is important that the selection of respondents to intercept studies continue to represent the people typically found using SRSNS. To date

[3] This is based on the assumption that a mean rating of 3.0 for effect on enjoyment and continued use is the threshold for unacceptability. It may be that 1.5 or 2.0 is more appropriate.

the intercept studies have been conducted over a weekend, or perhaps a couple consecutive weekends. It is advisable to increase sampling throughout the season of use at specific SRSNSs. In addition, it is also important to be vigilant that interest groups not learn of the dates and places of the intercept surveys, since if they do it may result in a "call to action" that in effect results in stuffing the ballot box. If this situation occurs, the respondents would no longer represent the evaluations of "typical" SRSNS users and the results should be discounted.

5 Summary Conclusions

This paper has demonstrated how to identify when users of a SRSNS find a scenic impact to be Unreasonably Adverse, based on user ratings of (1) a photograph of the actual view, (2) a photosimulation of how the view will appear if the wind development is constructed, (3) how the wind development will effect enjoyment of their use of the SRSNS, and (4) whether the wind development will affect their continued use of the SRSNS. Specifically, these data can be used to:

1. Evaluate the reliability of the scenic value ratings.
2. Calculate the effect size of the potential scenic impact.
3. Determine the threshold of Unreasonable Adverse scenic impact based on an assumed level where the project's effect on enjoyment of the SRSNS is unacceptable.
4. Determine the threshold of Unreasonable Adverse scenic impact based on an assumed level where the project's effect on the continued use of the SRSNS is unacceptable.

The data from four intercept surveys were reanalyzed, and their results were found to be both valid and reliable. A web-based survey was also reviewed and the reliability of the responses was lower than for the intercept surveys. However, the web survey does provide an opportunity for a more diverse public to contribute to the scenic impact assessment process. It is recommended that the effectiveness of using web-based surveys be further evaluated.

This paper is intended to initiate a discussion on how to identify the threshold between Adverse and Unreasonable Adverse scenic impacts. While the results appear impressive, they are based on relatively few data points. Future applications for wind energy development should include intercept surveys from more viewpoints that provide a greater range in the scope and scale of visible generation facilities, and are frequented by people engaged in a greater diversity of activities. Methods also need to be developed and validated to evaluate the effect of multiple exposures to scenic impacts while using SRSNSs. Intercept surveys should also be conducted in a manner that provides for estimating the extent, nature and duration of use at a viewpoint. Web surveys might be used to supplement intercept studies at locations with very few users, where there is the expectation of great controversy, or other situations where the scenic impact evaluation would benefit from a greater number of responses from potential users.

Finally, it is recommended that post-construction studies be conducted to monitor the actual scenic impact. In particular, intercept surveys should be conducted at the same locations as reviewed in this paper and include an evaluation of the perceived veracity of the photosimulations as a tool for evaluating scenic impacts.

References

American Psychological Association (2010), *Publication Manual of the American Psychological Association* (6th ed.). Washington, DC, APA.

Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). New York, Psychology Press.

Expedited Permitting of Grid-Scale Wind Energy Development (2008), MRSA Title 35-A, Chapter 34-A. http://www.mainelegislature.org/legis/statutes/35-A/title35-Ach34-A.pdf (accessed February 23, 2010).

Mildner, C. & MacBride, J. (2010), *Research Report: Mt. Blue-Saddleback Ridge Wind Power Project Intercepts*. Portland, ME, Market Decisions.

Mildner, C. & MacBride, J. (2010), *Spruce Mountain Wind Project Intercepts Research Report*. Portland, ME, Market Decisions.

Nunnally, J. C. (1978), *Psychometric Theory* (2nd ed.). New York, McGraw-Hill.

Palmer, J. F. & Hoffmann R. E. (2001), *Rating reliability and representation validity in scenic landscape assessments*. In: Landscape and Urban Planning, 54 (1-4), 149-161.

Portland Research Group (2011), *Highland Wind Project Survey of Hikers March 11: Report of Trail Intercept and Web Survey Research Conducted Summer and Fall 2010.* Portland, ME, Portland Research Group.

Robertson, B. & MacBride, J. (2010), *Research Report: Bull Hill Wind Power Project Intercepts*. Portland, ME, Market Decisions.

Roth, M. (2006), *Validating the use of Internet survey techniques in visual landscape assessment – An empirical study from Germany*. In: Landscape and Urban Planning 78, 179-192.

Stamps, A. E. (2000), *Psychology and the Aesthetics of the Built Environment*. Boston, Kluwer Academic Publishers.

Wherrett, J. R. (1999), *Issues in using the Internet as a medium for landscape preference research*. In: Landscape and Urban Planning, 45, 209-217.