Semantic Assessments of Experienced Biodiversity from Photographs and On-Site Observations – A Comparison

Mats Gyllin¹ & Patrik Grahn¹

¹ Department of Work Science, Business Economics and Environmental Psychology, Swedish University of Agricultural Sciences, Alnarp, Sweden

Correspondence: Mats Gyllin, Department of Work Science, Business Economics and Environmental Psychology, P.O. Box 88, SE-230 53 Alnarp, Sweden. Tel: 46-40-415-537. E-mail: mats.gyllin@slu.se

Received: June 27, 2015   Accepted: July 28, 2015   Online Published: October 27, 2015
doi:10.5539/enrr.v5n4p46          URL: http://dx.doi.org/10.5539/enrr.v5n4p46

Abstract
Since the 1960’s, public assessments of landscapes have often been carried out using photographic representations. How reliable and valid are these assessments compared with on-site observations? In the present study, participants have been asked to judge different areas in terms of a limited feature: the biodiversity of the area.

Digitalized photos from six different study areas were made available on the Internet, along with a questionnaire consisting of a semantic form with specific words/expressions to be rated in relation to the photos (four per area). Participants were recruited via mailing lists and informal contacts. These results were compared with a study in which students and ecologists had rated the same places using the same form, but this time on-site. The Internet participants were also asked to state their profession/education to make comparisons possible. The comparisons revealed differences between on-site and photo-based ratings, but the main difference was expressed by on-site biologists regarding areas with the highest experienced biodiversity values, possibly due to their higher degree of expertise and use of more senses than can be used when judging photographs. Concerning laymen in particular, it is concluded that the comparison between on-site and photo-based ratings is not conclusive enough to allow us to determine whether it is appropriate to use one method as a substitute for the other.

Keywords: biological diversity, perception, preference, public participation, semantic

1. Introduction
In 2007, for the first time in history, more people were living in cities than in the countryside (Obaid, 2007). In Sweden, cities with more than 100,000 people are growing particularly fast (Statistics Sweden, 2013). This rapid growth makes heavy demands on city planning. Among architects and city planners, a dense city is often considered to be a more suitable direction for future city planning than a sprawled city with regard to ideas of sustainable development: That is, a dense city involves fewer and shorter traffic movements (Young, 1995). The solution is often to densify existing urban environments, i.e. to build on the city’s green spaces. As a consequence, during the past 30 years, there has been an increasing tendency towards viewing urban green spaces as land reserves (Statistics Sweden, 2002, 2010).

1.1 Urban Sustainability - Biodiversity and Health
However, there are at least two other ways of defining sustainability: One idea that has political influence is the concept of biodiversity; the other is the idea of promoting people’s well-being and health. Research shows that urban green areas are important to the biodiversity of a region (Zerbe, Maurer, Schmitz, & Sukopp, 2003). Moreover, the benefit of green environments to human well-being has been demonstrated by many authors, at least since the 1980’s (Maas et al., 2009; Nilsson et al., 2011; Nutsford, Pearson, & Kingham, 2013). One way of explaining how urban nature can affect people’s health is through the biophilia hypothesis (Adevi & Grahn, 2011; Falk & Balling, 2009; Gullone, 2000; Kellert & Wilson, 1993), which attempts to explain our preference for, or even affection towards, other living beings using genetic factors. Hence, biodiversity and people’s health and well-being can be connected; e.g. by raising people’s level of physical activity, given that people prefer to visit urban environments with a large number of different species. This hypothesis is partly supported by several studies showing that natural environments are preferred to non-natural environments, particularly in urban settings (Hartig, Mang, & Evans, 1991; Herzog, 1989; van den Berg, Hartig, & Staats, 2007). Björk et al. (2008) also found...
a connection between exercise, BMI (body mass index) and a large number of qualities in the green environments close to where people lived, where biodiversity was mentioned as an important quality. Moreover, Björk et al. (2008) found that characteristics in nature that were not primarily visual were connected to health, which indicates that information taken in via senses other than sight are of utmost value to the public. This finding is to some extent confirmed by Diette (2003) and Annerstedt et al. (2013), both of which showed that sounds (birdsong) - together with the visual information in nature - was of great importance to people's preferences and their health. Another hypothesis is that biodiversity in green urban areas can promote people’s mental restoration owing to the feelings of fascination that occur in such environments (S. Kaplan, 1990; Ottosson & Grahn, 2005).

1.2 Defining Biodiversity

Studies concerning biodiversity and health (e.g. Adevi & Grahn, 2011; S. Kaplan, 1990), all relate to how people perceive biodiversity. How is this to be understood, and associated with how biodiversity usually is defined?

Biodiversity is an important matter, and has been considered ever since prehistory (e.g. Noah’s Ark and similar narratives in religious scriptures). In spite of its importance, during the last decades biodiversity has been handled and managed primarily in a technical way (Gyllin, 2004). Consequently, general confusion exists regarding the concept of biodiversity, because it can be viewed as either a scientific (biological) parameter (as in Hawkinsworth, 1995) or as a political tool aimed at nature conservation (as in the Convention on Biological Diversity, UNEP, 1992). Ehrlich and Ehrlich (Ehrich & Ehrlich, 1992), suggested four perspectives which could be applied on biodiversity: an ethic perspective (concerning responsibility of protecting different species from extermination); an aesthetic perspective (concerning the beauty of nature and its species, involving ecotourism, bird watching, scuba diving, nature films etc.); an economic perspective (concerning food-production, medicine or industrial products); and an ecosystem perspective (concerning ecosystem services, how an environment rich in biodiversity can control the gaseous mix of the atmosphere, maintain rich soils, and manage to control pests etc.). Luckett (2004), however, introduced two sets of perspectives. First, a set including six perspectives: Anthropocentric, animal-centric, species-centric, bio-centric, ecosystem-centric and eco-centric – from a pure Homo sapiens perspective to the whole biosphere. Next a set including two perspectives: a teleological perspective (that certain sorts of acts are right or wrong, based on an assessment of their consequences) and a deontological perspective (that certain sorts of acts are right or wrong themselves). To be more nuanced in debating biodiversity, Luckett (2004) claims that the two sets should be put together. For example, a teleological anthropocentric perspective could involve resource conservation and the idea of sustainable agriculture, while a deontological species-centric perspective could involve species preservation. In a study in Vanuatu, South Pacific, Caillon & Degeorges (2007) also found that people’s underlying foundation concerning preserving different species or not, was based on factors which involve many perspectives, not at least cultural factors, such as ancestral links. And in a study of urban biodiversity in a Turkish city, including ornamental plants, Acar, Acar, and Eroğlu (2007) found four clusters of plants. These clusters correlated significantly with the actual cultural variation in the city.

To most people vaguely familiar with the concept biodiversity, it involves a general sense of biological quality and/or quantity (Gyllin, 2004). Biodiversity in itself is supposed to reflect the variation of living creatures in a certain area, either as a measurement of certain indicator factors, or as an intuitive estimation, based on visual impressions and experience, resulting in values such as “high” or “low”. Hence, biodiversity can be approached in many different ways and at various levels of ambition (Gyllin, 2004), although it seems to be difficult to measure, particularly in its most general form.

1.3 How to Measure Perceived Biodiversity

In several studies (e.g. Hartig & Staats, 2006) perceived naturalness is presented as a visual property of the scenery, but the question is whether the preference pattern can be taken a step further, to include actual biological data, i.e. biodiversity.

One method of capturing public opinions in a structured way, using a semantic test, was presented by Gyllin and Grahn (2005). Their method enables identification of the perceived components of experienced biodiversity, and creation of a biodiversity experience index (BEI) to be used as a measurement of perceived biodiversity. Unlike the study by e.g. Caillon and Degeorges (2007), the index developed by Gyllin and Grahn (2005) is supposed to avoid ethical, aesthetical, economical and cultural perspectives on biodiversity (Acar et al., 2007; Ehrlich & Ehrlich, 1992). Hypothetically, the index is robust enough to bridge the gaps between different paradigms concerning the foundation of biodiversity.

Conducting such investigations is difficult, however, because they require the presence of a preferably large number of people. Although semantic methods are often considered highly reliable even when the number of
subjects is low, at least as low as 14-15 (R. Bruce Hull & Buhyoff, 1984; Küller, 1972, 1975), it would be far easier if the same study could be carried out using photographs of the areas instead.

1.4 Measuring Perceived Qualities in the Landscape through Photographs

Techniques for assessment of qualities in urban and rural green environments have been developed since the early 1900’s (Addison, 1931), and particularly since the 1960’s (Linton, 1968). These methods were often criticized for having more or less arbitrary value judgements, the rationale for which was not made explicit. Moreover, the techniques were based on expert judgements, and public opinion could have resulted in different priorities (Penning-Rosswell, 1975). Some attempts were made to involve public opinions, particularly in judging landscapes using photo-slides (Fines, 1968).

The most elaborated contribution in this area was made by Shafer and his associates (Shafer, Hamilton, & Schmidt, 1969; Shafer, 1964): the method known as “Scenic Beauty Estimations”. Since then, scenic beauty estimations have developed into an important issue in forestry, and have been used as a basis for people’s experiences (Buhyoff, Hull, Lien, & Cordell, 1986; Daniel & Boster, 1976; Daniel & Schroeder, 1979; Deng, Yan, Guan, & Katoh, 2012; Kalidindi et al., 1997; Real, Arce, & Sabucedo, 2000; Zhang, Ling, & Da, 2012). Scenic beauty estimations are based on colour slides representing forest areas; the slide pictures are taken in a standardized manner and shown to people, who then rate their ‘scenic beauty’ on a (usually) 10-grade scale. The method has been developed and refined since the beginning of the 1970s and has reached a high level of reliability.

Photo-based questionnaires as a phenomenon do not decrease in our society - on the contrary. There is a will from authorities in many countries to involve the public more and more in decisions affecting the landscape. This applies to the construction or expansions of roads, dams, tourist facilities, etc., and its impact on, inter alia, noise and biodiversity. Such studies usually include questionnaires based on photos (Barroso, Pinto-Correia, Ramos, Surowá, & Menezes, 2012; Pinto-Correia et al., 2013; Sherren et al., 2011; Tveit, 2009).

Today, photo-based questionnaires are not as often contested as a couple of decades ago. However, the interest starts to rise again (Le Lay, Cottet, Piégay, & Rivière-Honegger, 2012; Sevenant & Antrop, 2011). Still, the question is whether these methods are valid in the sense that photo-based assessments can be considered equivalent to evaluations made on-site.

There has been some controversy in the scientific literature regarding the usefulness of pictures in perception studies on outdoor environments, particularly regarding the following aspects that illustrate flaws in either on-site studies or in studies on photos:

- Flaws in on-site investigations depending on participants’ physical efforts, such as bad weather, long distances to walk or cover by other means of transportation and lengthy periods of time that may cause weariness. Hull and Stewart (R. B. Hull & Stewart, 1992) found that judgements based on photos differed from on-site judgements, if the latter involved some physical effort and a gradually changing trail along which the judged areas were scattered. Furthermore, their results indicated that there is a risk for systematic errors that may bias group averages, as they found substantial differences in individual judgements, which were cancelled out by group averages.

- Flaws in on-site studies, depending on too much inappropriate information. Judgements based on photos are often considered equivalent to on-site experience (García Pérez, 2002; Hartig, Korpela, Evans, & Gärling, 1996; Hultman, 1983; R. Kaplan & Kaplan, 1989, pp. 15–17), or even superior. Some authors (García Pérez, 2002; Hultman, 1983) have argued that the context of on-site judgements may actually disturb the experience. In general, photo-based investigations seem to be regarded favourably by most of the researchers mentioned above, although they sometimes admit that they are not the same as on-site investigations (e.g. Hultman, 1983). The question is whether they are sufficiently similar for the purposes of the particular investigation.

- Flaws in photo-based studies, depending on the duality of photographic representation: Is it a representation of a place or is it a piece of art? Scott and Canter (Scott & Canter, 1997) noted that people familiar with an area shown in a photo responded differently depending on whether they were asked to comment on the place or the picture. Thus, there is a duality in a photograph (or any picture), because it is an object in its own right at the same time as it is a representation of something else (Gibson, 1986, p. 280 ff.).

- Flaws in photo-based studies, depending on lack of appropriate information. Brown et al. (1989) registered substantial differences between photo-based and on-site preference if the participants had actively chosen the judged areas as a campground. However, the authors did not offer any definite explanations for this. Nor did they take sides as to which method is preferable. Gibson’s (1986) ecological approach to vision might explain some of the differences, apart from the duality in viewing photos mentioned above. The core of Gibson’s
approach implies that perception in an environment is largely the result of movement (ambulatory vision) and the ability to look at objects from different angles (ambient vision). According to Gibson, environmental perception is a very dynamic experience, as opposed to the static activity of viewing the ‘frozen’ environment in a photo. Thus, photos can only display a very limited amount of the unlimited information provided by the ‘live’ environment.

- Flaws in both photo-based and on-site studies, depending on the choice of informants. In some cases, judgements differ depending on who is making the judgements, particularly if there are ‘experts’ in some sense involved. van den Berg et al. (van den Berg, Vlek, & Coeterier, 1998) made an interesting attempt to interpret differences among different groups of subjects in their beauty experience of areas that were also assigned a certain biodiversity value, either by the subjects themselves or by experts. Their results indicate that beauty ratings for environments with high levels of perceived biodiversity are also high for all investigated groups of subjects. However, when experts in nature conservation estimated biodiversity, beauty judgements differed between groups. Particularly the judgements of farmers resulted in low beauty estimates for high biodiversity values and vice versa. The conclusion relevant to the present paper is that farmers and nature conservationists have different opinions about what biodiversity is. We suggest, therefore, that biodiversity ratings are to some extent aesthetic.

The main aims of the present study have been
- to investigate whether the judgement of perceived biodiversity differs if made from a picture or on-site, and
- to investigate whether the judgement differs if it is made by biologists or non-biologists.

We hypothesized that
- ratings of biodiversity quality based on photos would differ from on-site judgements, based on the assumption that available information differs, and that
- ratings of biodiversity quality made by biologists would differ from judgements made by non-biologists, based on the assumption that the ability to interpret the perceived environment differs.

2. Methods

2.1 Protocol

To measure the content of experienced biodiversity in urban green areas, we used the Biodiversity Experience Index protocol, BEI (Gyllin & Grahn, 2005). Semantic environment description as a method briefly means that each member of the test panel estimates – according to a Likert scale (Barnett, 1991) - to what extent a number of words fit with a particular environment. Description of environments through semantic studies has been established for several decades as a tool to survey and measure experience categories in various built environments, and estimations by a total of 15-20 persons is considered reliable enough to reach stable results (Küller, 1972, 1975). Since the 1970’s, methods have been established to measure experience values in landscape architecture (Axelsson Lindgren, 1990, 1999; Grahn, 1991) and some studies also include elements that can be directly connected to species richness (Berggren-Bärring & Grahn, 1995; Caspersen & Olafsson, 2006; Grahn, Stigsdotter, & Berggren-Bärring, 2005; Grahn & Stigsdotter, 2010; Stockholms läns landsting, 2001).

Words that could be potential contributors to the expression ‘biological diversity’ were selected and included in the study by Gyllin and Grahn (2005). A form was created with a scale at each word from 1 (“A little”) to 5 (“Much”). All these words were tested in different semantic studies to high reliability and validity; not least concerning codability, i.e. they were all understandable in the everyday language (Furberg, 1982; Gyllin & Grahn, 2005, 2005; Sigurd, 1983, 2002). The final protocol included four words.

Our aim is to make the protocol valid and reliable. Hence, as in Gyllin and Grahn (2005) the BEI for participant p’s rating of study area s was calculated using a formula for weighted averages (Ter Braak & Barendregt, 1986). A Varimax factor analysis (Manly, 1994) resulted in factor loadings used in the following formula:

$$BEI_{ps} = \frac{\sum_{i=1}^{n} (FL_i \cdot R_i)}{\sum_{i=1}^{n} FL_i}$$

$FL_i$ is the factor loading for the $i^{th}$ of the used words/expressions in the instrument. $R_i$ is the rating value of the $i^{th}$ used word/expression. $n$ is the number of used words/expressions, which in the case of BEI is four: ‘Plant richness’, ‘Animal richness’, ‘Wild’, and ‘Varied’ (Gyllin & Grahn, 2005).
2.2 Study Areas

The study was carried out in six different areas, representing a stratified selection of common southern Swedish park types with varying management intensity. They were arranged in three pairs, each pair internally contrasting and representing a particular theme. All areas were 1-2 hectares, a common size for green spaces in the region’s urban areas. Both areas of each pair were within walking distance from each other. The study areas are all in southern Sweden (the cities of Malmö, Lund, and Eslöv).

Theme: Green areas with water in the urban fringe. These areas are situated on the outskirts of Malmö and may be characterized as wetland parks. Parks with water should potentially contain more species, and the hypothesis is that these areas should obtain higher biodiversity estimates.

Area 1 (Malmö): The Toftanäs marl-pit. A relatively simple pond used to take out calcareous soil (’marl’) as fertilizer. The vegetation along the edges of the water is dense and free growing, while the rest of the area consists of simple lawn with a few planted trees and some artificial hills. Figure 1 shows the pictures from Area 1 used in the study.

Figure 1. Area 1 – “The Toftanäs marl-pit” in the eastern outskirts of Malmö

Area 2 (Malmö): The Toftanäs wetland park. A relatively simple pond used to take out calcareous soil (’marl’) as fertilizer. The vegetation along the edges of the water is dense and free growing, while the rest of the area consists of simple lawn with a few planted trees and some artificial hills. Figure 2 shows the pictures from Area 2 used in the study.

Figure 2. Area 2 – “The Toftanäs wetland park” in the eastern outskirts of Malmö
Area 2: The Toftanäs wetland park. A storm water detention facility designed to imitate a natural wetland with a meandering watercourse linking two small ponds together. The number of plant species has been increased by introducing meadow species of local origin. Also the spontaneously invading flora contributes to increased species richness, primarily from a ditch leading into the area, but also wind-dispersed species and species appearing in the seed bank. The facility is described in detail in Gyllin (2004). Figure 2 shows the pictures from Area 2 used in the study.

Theme: Central, rectangular green areas without water. Two areas from central Lund that were part of the investigation reported in Berggren-Bärring and Grahn (1995), where they represented two extremes regarding words connected to biodiversity.

Area 3: Holmgrenska tomten. Previously privately owned area, now owned by the municipality of Lund. It can be characterized as a small elm forest, partly very dense and brushy. In a corner of the area where there have been allotment gardens, the trees are younger and even denser. Figure 3 shows the pictures from Area 3 used in the study.

Figure 3. Area 3 – ”Holmgrenska tomten” in the central parts of Lund

Figure 4. Area 4 – ”Mästers park” in the central parts of Lund
Area 4: Mästers park. Simply designed park with a short-cut lawn surrounded by a hedge-like shrubbery of mostly Symphoricarpos rivularis and a single row of mature lime trees. A well-trodden, spontaneous footpath leading diagonally across the lawn, suggests that the park is used as a short-cut. In a section of the park, there is a small hornbeam stand with a playground. Figure 4 shows the pictures from Area 4 used in the study.

Theme: Central, linear green areas without water. Two areas from Eslöv representing linear or corridor elements. They differ mainly in management, where Area 5 is rather strict with a few tree species (mostly poplars) and most of the area covered with short-cut grass. Area 6 is distinctly ‘wilder’ with spontaneous vegetation.

Area 5: The narrow lawn park. A long and narrow area, chiefly consisting of an ordinary short-cut lawn. On one side it borders to a sports field, belonging to a nearby school. On the other side it borders to a residential area. Close to the sports-field side there is a spontaneous footpath through the entire length of the area. Figure 5 shows the pictures from Area 5 used in the study.

Figure 5. Area 5 – “The narrow lawn park” in the northern parts of Eslöv, close to the Norrevång school

Figure 6. Area 6 – The railway embankment on the Eslöv – Landskrona route, close to area 5
Area 6: The railway embankment. This surrounds a portion of the railway between Eslöv and the town of Landskrona, only a few tens of meters from the end of Area 5. It is fenced, on one side bordering to a public green open space (lawn) and on the other side to a residential area. Spontaneously appearing trees and bushes make it difficult to observe. Figure 6 shows the pictures from Area 6 used in the study.

Based on an inventory of the different sites (concerning species-richness – plants, birds, insects), we expected area 2 to get the highest BEI-assessment followed by area 3, area 1, area 6 and area 5, ranked in order. Area 4 was expected to have the lowest BEI-assessment.

2.3 Test Groups

The group judging the areas on-site consisted of 102 persons (70 women and 32 men). They were brought (by bus) to the six different urban areas chosen for this study, and were asked to complete the Biodiversity Experience Index form. The participants were recruited among first-year students at the Swedish University of Agricultural Sciences in Alnarp, Sweden, plus senior researchers in plant ecology at the University of Lund. The number of students was 96 and of senior researchers six. The average age of the participants was 26 years, ranging from 18 to 54.

To obtain data from photo-based judgements, digital photos were made available on the Internet, along with instructions, as a World Wide Web site. Web sites as a medium for enquiries have been in use since the early 1990’s, and have been found to be reliable, however difficult to control regarding randomness and how well they represent the “off-line” population (Andrews, Nonnecke, & Precece, 2003). Participants were recruited through mailing lists and informal contacts to a final number of 124 (57 women and 67 men). The average age was 38, ranging from 14 to 63.

2.4 The Online Interface

The electronic form was as near identical to the paper form used in the on-site study as possible. Apart from the Biodiversity Experience Index protocol, participants were asked to provide information on their age, gender, education and professional background, just as in the paper form, but they were also asked to state whether they had participated in the on-site investigation (there was only one such case). Based on professional background, it was possible to determine which of the participants were biologists. The number of participants who claimed that their background or profession was as biologists (or ecologists, limnologists, or other biological specialists) were 36 and non-biologists were 88. The backgrounds of the non-biologists varied, but an overwhelming majority had some university level education. Attached to each word were five ‘radio buttons’, one for each grade of the scale. It was only possible to select one of them, but it was possible to avoid selecting any of them. The objects to be judged were the same six areas used in the on-site study, represented by six photographs each. Thus, the entire set consisted of 36 photos (Figures 1–6), which were all taken by the first author in full daylight on May 16. For each area, all six photos were initially seen simultaneously as miniatures on the screen, but each one of them could be enlarged separately to cover the entire screen. At the bottom of the screen, the form was then submitted to the first author by email.

2.5 Statistics

Differences in BEI among study areas were investigated using repeated-measures analysis of variance (ANOVA), because the ratings for all areas were made by the same persons. One group in the study has a quite small sample, which might cause a problem concerning the ANOVA assumption of variance homogeneity. To check for this, Levene's test was used.

Differences between groups of participants and between on-site and photo studies were investigated using two-way ANOVA (Sokal & Rohlf, 1995).

Differences between groups of participants for all sites separately for the on-site and photo studies were investigated using two-sample t-tests (Sokal & Rohlf, 1995). F-test was used to test for equal variances.

Statistical analyses were made in Statistica, version 5.5 for Windows (StatSoft Inc., 1995) and Minitab Statistical Software (Minitab Inc., 2006).

3. Results

From the factor analysis, Varimax rotated, with all the 226 enquiries, the factor loadings of the four used words/expressions from factor 1 are presented in Table 1.
Table 1. Results from factor analysis, Varimax rotation, based on all 226 enquiries. Only loadings from the used words/expressions are shown

<table>
<thead>
<tr>
<th>Word/expression</th>
<th>Factor loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plant richness</td>
<td>0.79</td>
</tr>
<tr>
<td>Wild</td>
<td>0.75</td>
</tr>
<tr>
<td>Animal richness</td>
<td>0.71</td>
</tr>
<tr>
<td>Varied</td>
<td>0.55</td>
</tr>
</tbody>
</table>

The Biodiversity Experience Index, BEI, was calculated from weighted variables, using the formula for weighted averages. Differences in BEI among study areas were highly significant, according to the repeated measures ANOVA, except between Area 2 and 3. Means and 95% confidence limits are presented in Figure 7. Levene’s test for homogeneity of variances showed no effect.

The further analyses of differences in BEI between on-site and photo-based ratings as a two-way analysis of variance using the category variables study (on-site, photo-based) and occupation (biologist, non-biologist) revealed varying degrees of significant differences, summarized in Figure 8–13. For Area 1 (the Toftanäs marl-pit, Figure 8), there is just a weak separation of on-site ratings from the photo-based. Area 2 (the Toftanäs wetland park, Figure 9) and 3 (Holmgren’s tomten, Figure 10) show similar patterns, although in Area 2, the ratings of on-site experts are higher than the others, while in Area 3, the ratings of on-site non-experts are lower than the others. Area 4 (Mästers park, Figure 11) and 5 (the narrow lawn park, Figure 12) also show similar patterns, because experts and non-experts are separated, as well as on-site and photo-based studies, but there is no interaction effect, i.e. experts and non-experts are not separated by study type. Area 6 (the railway embankment, Figure 13), finally, shows a weak separation of on-site non-experts from photo-based experts.
Figure 8. Results from the two-way ANOVA for Area 1 showed a significant difference between the on-site and photo-based studies ($F=6.1$, $p<0.05$), but no difference between experts and non-experts, and no interaction effect.

Figure 9. Results from the two-way ANOVA for Area 2 showed a significant interaction effect ($F=13.9$, $p<0.001$). Post hoc analysis (Tukey HSD) revealed differences between experts on-site and all the other groups. The one-way comparisons were also significant ($p<0.05$ for on-site/photo, and $p<0.01$ for experts/non-experts).

Figure 10. Results from the two-way ANOVA for Area 3 showed a significant interaction effect ($F=15.9$, $p<0.001$). Post hoc analysis (Tukey HSD) revealed differences between non-experts on-site and all the other groups. The one-way comparison between experts and non-experts was also significant ($p<0.01$).

Figure 11. Results from the two-way ANOVA for Area 4 showed a significant difference between the on-site and photo studies ($F=5.7$, $p<0.05$), as well as between experts and non-experts ($F=8.1$, $p<0.01$), but no interaction effect.

Figure 12. Results from the two-way ANOVA for Area 5 showed a significant difference between the on-site and photo studies ($F=11.3$, $p<0.001$), as well as between experts and non-experts ($F=6.2$, $p<0.05$), but no interaction effect.

Figure 13. Results from the two-way ANOVA for Area 6 showed a significant interaction effect ($F=4.6$, $p<0.05$). Post hoc analysis (Tukey HSD) only revealed a difference between non-experts on-site and photo-based experts. The one-way comparison between on-site and photo was also significant ($p<0.05$).
A t-test, performed to summarize differences in BEI data between biologists and non-biologists, is presented in Table 2. The table shows differences between the groups, but not in the same way on-site as in the photo-based study. Only area 5 showed significant differences in both studies.

Table 2. Summary table of two-sample t-tests (separately for each study and each area) of differences in BEI between biologists and non-biologists. Significant p-values mean that the area was rated differently by biologists than by non-biologists. Mean diff are differences in the mean values, where positive values indicate a higher mean value for biologists and negative values for non-biologists. NS (not significant) represents p-values larger than 0.05. On-site biologists were 6, non-biologists 96. Photo-based biologists were 36 and non-biologists 88. T is the t-test T value, and DF represents the degrees of freedom. Areas 2 and 5 on-site showed unequal variances (F-test, p<0.05), hence the reduced DF

<table>
<thead>
<tr>
<th>Area</th>
<th>On-site</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean diff</td>
<td>T</td>
<td>DF</td>
<td>p</td>
<td></td>
<td>Mean diff</td>
<td>T</td>
<td>DF</td>
<td>p</td>
<td></td>
<td>Mean diff</td>
<td>T</td>
<td>DF</td>
<td>p</td>
<td></td>
<td>Mean diff</td>
<td>T</td>
<td>DF</td>
<td>p</td>
<td></td>
<td>Mean diff</td>
<td>T</td>
<td>DF</td>
<td>p</td>
</tr>
<tr>
<td>1 (Toftanäs marl-pit)</td>
<td>0.241</td>
<td>0.74</td>
<td>99</td>
<td>0.463 (NS)</td>
<td></td>
<td>-0.093</td>
<td>0.65</td>
<td>122</td>
<td>0.515 (NS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 (Toftanäs wetland p.)</td>
<td>1.006</td>
<td>13.39</td>
<td>51</td>
<td>0.000 (*** )</td>
<td></td>
<td>-0.078</td>
<td>0.62</td>
<td>122</td>
<td>0.534 (NS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 (Holmgrenska tomten)</td>
<td>1.131</td>
<td>3.67</td>
<td>100</td>
<td>0.000 (*** )</td>
<td></td>
<td>-0.146</td>
<td>1.13</td>
<td>122</td>
<td>0.259 (NS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 (Mästers park)</td>
<td>-0.543</td>
<td>-2.70</td>
<td>100</td>
<td>0.008 (** )</td>
<td></td>
<td>-0.195</td>
<td>-1.61</td>
<td>122</td>
<td>0.110 (NS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 (The narrow lawn p.)</td>
<td>-0.450</td>
<td>-3.73</td>
<td>10</td>
<td>0.004 (** )</td>
<td></td>
<td>-0.343</td>
<td>-2.36</td>
<td>121</td>
<td>0.020 (*)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 (The railway emb.)</td>
<td>0.401</td>
<td>1.20</td>
<td>99</td>
<td>0.232 (NS)</td>
<td></td>
<td>-0.326</td>
<td>-2.46</td>
<td>120</td>
<td>0.015 (*)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Discussion

In this study, our aim was primarily to investigate whether the judgement of perceived biodiversity differs if made from a picture or on-site, and secondarily to investigate whether the judgement differs if made by biologists or non-biologists.

The present investigation shows the robustness of BEI when applied to photos as compared to on-site assessments. Figure 7 shows clear and in many cases significant differences between the areas, and Figures 8-13 show that on-site and photo-based assessments are similar, but not identical. Assessments for all areas differ slightly, but significantly. Some of these differences may be explained by the difficulty of taking representative photos; some may depend on the selection of participants, though others may principally be explained by the fact that photos simply cannot provide enough information.

The total BEI assessment, showed in figure 7, is quite close to what we had expected, based on our inventory of the different sites. This is an interesting finding, and should be viewed in the light of articles focusing different biodiversity paradigms, may it be cultural, moral or scientific (Ehrich & Ehrlich, 1992; Luckett, 2004). In the BEI protocol, words that are found to be easy to interpret and understand by the general public have been used (Furberg, 1982; Gyllin & Grahn, 2005; Sigurd, 2002). So cautiously, this can be understood by the simplicity of the instrument.

As pointed out earlier, using photos to assess environments is far easier than making such assessments in the field. The question is whether such an assessment is good enough, or perhaps more accurately, whether it achieves what it is expected to. Advocates of photo-based as well as on-site studies have mentioned an array of confounders:

- Flaws in on-site investigations due to the participants’ physical efforts could affect judgements. There were differences between on-site and photo-based BEI ratings, sometimes when on-site non-experts gave lower ratings than in the photo study, and sometimes when on-site experts gave higher ratings than in the photo study. Hence, nothing indicates that physical effort was a significant factor, especially as the last area visited (Area 6) differed very little from the others. There were no comments regarding this either. The effects of effort demonstrated by Hull and Stewart (R. B. Hull & Stewart, 1992) probably emerged because the exertions in their study were more severe.

- Flaws in photo-based studies due to duality in representation could induce uncertainty as to what to judge. Just as Gibson (1986) pointed out, there is an ambiguity in photos that could ‘trick’ the observer into experiencing more beauty from the picture than from an actual visit – the photo as a work of art. The results in Table 2 indicate that the six on-site ecologists mainly differ from the non-experts concerning Area 2, 3, 4 and 5. Particularly Area 2 (the Toftanäs Wetland Park) and 3 (Holmgrenska tomten) were rated substantially higher by the experts. Those two areas
are very unique environments for urban areas, Area 2 being a wetland created to resemble ‘nature’, and Area 3 a stand of ‘wild’ looking woodland in the central parts of Lund. It seems likely that the trained senses of skilled plant ecologists find things that non-experts do not and that are not easily detected in the photos. One of the reasons why the non-experts rated Area 3 low in the on-site study may be that the forest floor was very dry during the investigation period. This may have affected their judgements, while the experts probably emphasized the potential of the area, rather than the temporary drought. Smells, sounds, and subtle signs may trigger such expert abilities. The forest floor was not clearly visible in the photos, and the photo-based ratings followed the on-site experts, so it is possible, however uncertain, that the drought induced the significant difference shown in Figure 10.

The same is not true for the photo-based study, where biologists seem to differ less from non-experts, and in a more complex manner. How this should be interpreted is difficult to say, but perhaps some of the explanation simply lies in the fact that the biologists and non-biologists in the photo-based study were less different from each other than in the on-site study. In combination with a more standardized source of information (photos), personal preferences may be relatively more important than professional background.

Flaws in photo-based ratings that depend on the differing amount of information could affect how the environments are perceived, through lack of appropriate information as well as too much inappropriate information. Sight is a sense of utmost importance when visiting urban green spaces. However, other senses are also of great importance. The context of the presented environments is not visible in the photos, i.e. surrounding streets, signs of use or other visible phenomena that often affect the way we perceive the urban environment (again, see Gibson, 1986). In addition, impressions processed through senses other than vision are lacking: temperature, noise, moisture, odours, etc. Whether this is good or bad depends on the purpose, as Brown et al. (1989) concluded, but in the case of assessing biodiversity, the smell of flowers, fungi and decay may have an effect. Also the sounds of birds or insects and the feeling of moisture could affect assessments of biodiversity. Research results indicate that information derived from our senses is coded and stored in three different ways: subymbolic, symbolic imagery, and symbolic verbal (Bucci, 2000). Subsymbolic processing operates in sensory, motoric, and somatic modalities, which help, e.g., the wine taster judge the quality of the wine and the physician analyse a patient’s internal state by using, e.g., his/her nose and hands. All these knowledge processes occur in specific sensory-somatic modalities rather than in systematic cognitive ways: Information is processed via muscles, our inner organs, etc. (Bucci, 2000). The skilled biologist presumably uses subymbolic information just as the wine taster or the physician does. In Table 2 we find that BEI-assessments differ significantly between experts and laymen as regards five of the six sites in the study. However, we find the biggest differences between experts and laypersons pertaining to on-site ratings in area two and three – the areas we, based on inventories, hypothesized would have the highest BEI-ratings. Here, the experts on-site gave considerably higher grades than laypersons.

A preliminary general conclusion could be that photo-based investigations concerning biodiversity are less suitable for expert judgements, whereas they may be more suitable for non-experts, as the judgements of non-experts may be more affected by arbitrary factors such as drought, physical effort, traffic noise, etc. This should be tested more thoroughly, however, in studies focused on that specific issue. Nilsson (1988) – who found substantial differences in judgements of industrial landscapes between experts (professors in landscape architecture) and students – concluded that one possible explanation for the different judgements of industrial landscapes between students and professors may be that the professors’ higher degree of professional self-confidence makes them ‘bolder’ in appreciating evidently designed industrial landscapes. Similarly, in this study, the on-site experts are skilled researchers who spend considerable time working in and interpreting wilderness areas, possibly making them ‘bolder’ in their ratings of the most wilderness-like areas, 2 and 3. According to Kaplan (1988), there is substantial evidence that experts perceive differently within their area of expertise. Kaplan regards this as an important facet of their expertise: they perceive differently because they are experts. In a study by Bonnes and colleagues (Bonnes, Uzzell, Carrus, & Kelay, 2007) it was also concluded that laymen (residents in the quarters being estimated) and experts in plant ecology rated satisfaction with the quality of urban green spaces in the quarters differently. The authors interpreted the results that usability of the urban green spaces was more important to the residents than the amount of biodiversity in the area. Biodiversity can be viewed as one dimension of usability, but there are of course others (Grahn et al., 2005). In the present study, we can clearly see the effect of expertise in the on-site study, while the experts in the photo-based study either have too limited information or are not experts enough. According to Palmer and Hoffman (Palmer & Hoffman, 2001), the information in single photographs as representations of landscapes is too limited, and although the areas studied here are represented by six photos each, the information may still be too limited.

The participants were not the same in the both studies, not recruited in the same way, and were different in terms of average age and gender proportions. It should be emphasized that the two groups of ‘experts’ differed somewhat.
In the Internet study, the ‘experts’ were those who claimed that they were ‘biologists’, which could mean different things, from teachers of biology to nature conservationists and researchers. The experts in the on-site study were hand-picked senior researchers, all of whom had published their doctoral theses in plant ecology at the same department, making them a rather exclusive and homogenous group. Thus, all conclusions should be made rather cautiously, but there are nevertheless interesting points to discuss. At first, we may state that because of their homogeneity, the on-site experts produced very reliable results, as there were actually statistically significant differences compared to the on-site non-experts (Figures 9-10). We plan to carry through studies with an aim to validate these findings, involving people more randomly selected; laymen as well as experts.

It is possible that this kind of study requires additional photos. In the Scenic Beauty Estimation method (Daniel & Boster, 1976; Daniel & Schroeder, 1979), photos were taken in standard positions, but it seems doubtful that this would result in good representations. A better way could be if more than one person took photos of the areas independently. The character of the areas tested was rather distinct in the sense that they could be considered pair-wise extremes regarding the expected effect on biodiversity ratings. Perhaps a choice of more complex environments would have made the results more complex, too. Thus, the test of robustness should continue with a wider selection of environments and with more groups of participants.

5. Conclusions

The following conclusions can be made based on the results discussed above:

Our interpretation of the results presented is that there is a definite tendency for experts in the field of biology to make deeper, subsymbolic assessments of biodiversity when making on-site judgments.

The results of the BEI assessments for laymen are not conclusive enough to allow us to determine whether they can be used to make photo-based assessments that are comparable to on-site assessments. More testing and refining is needed before this can be achieved.

Acknowledgements

This paper is based on a research project funded jointly by the Swedish Research Council Formas and the Swedish University of Agricultural Sciences.

The authors would like to express their gratitude to Jan-Erik Englund, PhD, at SLU, for his valuable comments and corrections regarding the statistical methods.

References


Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).