Ecological Indicators 57 (2015) 293–297

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Confirmation bias in studies of fluctuating asymmetry

Mikhail V. Kozlov*, Elena L. Zvereva

Section of Ecology, Department of Biology, University of Turku, 20014 Turku, Finland

ARTICLE INFO

Article history:
Received 28 January 2015
Received in revised form 29 April 2015
Accepted 4 May 2015

Keywords:
Betula pubescens
Environmental stress
Fluctuating asymmetry
Indication of stress
Industrial pollution
Research bias

ABSTRACT

Fluctuating asymmetry (FA) represents small, non-directional deviations from perfect symmetry in morphological characters. FA is generally accepted to increase in response to stress; therefore, FA is frequently used in ecological studies as an index of stress experienced by an organism, in particular due to environmental pollution. We experimentally tested the hypothesis that the outcomes of studies based on FA measurements may have been influenced by confirmation bias, i.e. the tendency of humans to seek out evidence in a manner that confirms their hypotheses and beliefs. We collected 100 leaves of downy birch (*Betula pubescens*) from a single tree, grouped them haphazardly into ten samples, scanned every sample, and then asked each of 31 scientists (experienced in studying FA) to measure FA from the scanned images of all 100 leaves. Three groups of participants were provided with false information about the origin of each sample and about the hypothesis to be tested, and one group (control) was provided with true information. The participants who believed that the leaves originated from a heavily polluted site reported significantly higher values of FA when compared to the participants who believed that the leaves were collected from an unpolluted site. When the participants were told that half the samples originated from a polluted site and half from a clean site and were asked to attribute each sample to either of these sites based on leaf FA, the differences in FA between samples classified as 'polluted' and 'unpolluted' were much higher than the differences obtained from measurements of the same sets of samples made by the control group. We conclude that when scientists expected to find high FA in some samples, the results of their measurements confirmed their expectations. This effect, classified as confirmation bias, may considerably influence the outcomes of the research on FA. This confirmation bias can be avoided by using a blind method, where the person conducting measurements is not aware of the origin of samples being measured. We argue that the use of blind methods is critically important for any study addressing environmental or genetic impacts on FA.

1. Introduction

Fluctuating asymmetry (FA) represents small, non-directional deviations from perfect symmetry in morphological characters. It is commonly accepted that FA results from the inability of an individual to control development while under genetic and/or environmental stress (Palmer and Strobeck, 1986; Møller and Swaddle, 1997; Leamy, 1999). The very first discoveries that FA increases in response to different disturbances led to numerous enthusiastic recommendations to use FA as a handy indicator of stress experienced by organisms (Zakharov, 1990; Clarke, 1992; Parsons, 1992; Freeman et al., 1993; Hume, 2001). More specifically, FA was advertised for evaluation of environmental health (Zakharov and Clarke, 1993), the impact of environmental pollution (Eeva et al., 2000), quality control in mass rearing of insects (Clarke and Mckenzie, 1992), monitoring animal welfare (Knierim et al., 2007) and identification of the extinction risk for threatened populations (Anciães and Marini, 2000). A number of studies did not detect any changes in FA in response to environmental and genetic stressors (for examples, consult Graham et al., 2010), and yet the idea that FA increases with stress became very popular among scientists. The seeming simplicity of the measurements of FA resulted in wide application of this method in ecological and environmental studies, and the number of confirmatory publications is increasing rapidly. Consistent with this trend, a recent meta-analysis (Beasley et al., 2013) promoted the use of FA as a biomarker of environmental stress, which is likely to further increase the use of FA as indicator of environmental quality (e.g., Erofeeva et al., 2011; Chudzinska et al., 2014; Klisarić et al., 2014; Shadrina and Volpert, 2014).

Various biases occur in the planning, data collection, analysis and publication phases of scientific research and these are known to have a substantial influence on its outcomes (Pannucci

* Corresponding author. Tel.: +358 2 3335716.
E-mail address: mikoz@utu.fi (M.V. Kozlov).

http://dx.doi.org/10.1016/j.ecolind.2015.05.014
1470-160X/

and Wilkins, 2010). Confirmation, observer or expectancy bias—the tendency of humans to seek out evidence and interpret it in a manner that confirms their existing ideas and hypotheses (Rosenthal, 1976; Nickerson, 1998)—is a well-documented phenomenon in psychology and cognitive science. This bias results primarily from automatic processes occurring unintentionally (Hergovich et al., 2010). In contrast to the widely appreciated influence of publication bias on the understanding of ecological processes (Jennions et al., 2013), the occurrence and importance of biases introduced at earlier (pre-publication) stages of ecological research have received little attention (but see Loehle, 1987; Kozlov and Zvereva, 2009; Koricheva et al., 2013). Within biological disciplines, confirmation bias was demonstrated to affect the outcomes of studies on animal behaviour (Marsh and Hanlon, 2007; van Wilgenburg and Elgar, 2013) and on insect herbivory (Kozlov et al., 2014).

In this study, we experimentally tested the hypothesis that the outcomes of studies based on non-blind FA measurements (when the person conducting measurements was aware of the origin of samples being measured) might have been influenced by confirmation bias. In particular, we addressed the possibility that the influential theory that pollution increases leaf FA in plants (Freeman et al., 1993; Zakharov et al., 2001) could bias the results of measurements in the direction that would favour that theory. We asked 31 volunteer scientists to measure FA in the same 100 leaves collected from one birch tree growing in a pristine forest, while providing the participants with either true or false information regarding the sample origin. We predicted that: (1) the participants who believed that the samples were collected from a heavily polluted area would report higher values of FA than would the participants who believed that the exact same samples were collected from an unpolluted area; and (2) when the participants believed that the samples originated from both polluted and unpolluted sites, the differences in the reported FA values between samples classified as 'polluted' and 'unpolluted' would be higher than when the participants knew that the exact same samples originated from one individual plant (i.e. no heterogeneity would be expected).

2. Material and methods

2.1. Samples and participants

About 200 leaves with no traces of mechanical damage, deformation or insect feeding were collected on July 21, 2014 from a single tree of *Betula pubescens* Ehrh. (Betulaceae) near Turku, Finland. This species was selected for our study because birches are commonly used for studying environmental impacts on leaf FA (Kryazheva et al., 1996; Valkama and Kozlov, 2001; Hagen et al., 2008; Kozlov et al., 2009; Erofeeva et al., 2011; Shadrina and Volpert, 2014). The leaves were pressed between the sheets of filter paper and dried as ordinary herbarium specimens. A selection of 100 perfectly preserved leaves was haphazardly divided into 10 samples, each sample was scanned with high resolution and the scale was added to each image.

The participants were recruited by sending e-mail invitations to 84 corresponding authors of studies addressing FA in plants that were published in 2009–2013; 31 scientists agreed to participate in the project (for the list of participants, consult Appendix A). As the authors of published studies, all these scientists were experienced in FA measurements. The participants were told that the goal of the study was to test the reproducibility of the measurements, i.e. the consistency between the results obtained in different laboratories, by different operators and by using different instruments.

The participants were asked to measure the width of the left and right halves (the character which is commonly used to quantify FA) of each of 100 leaves at the midpoint between the base and the apex of leaf lamina (WL and WR; Fig. 1) using the same instruments and methods as they had used in their earlier studies, and calculate FA as follows: $FA = 2 \times |WL - WR|/(WL + WR)$.

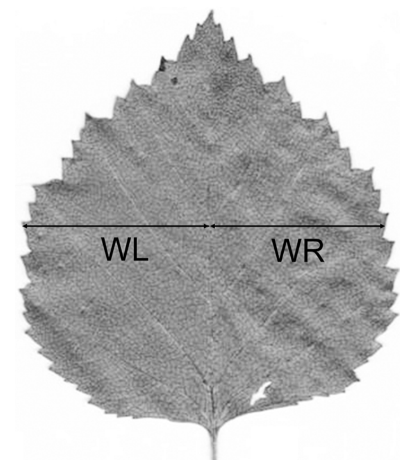

Fig. 1. Measurements of left (WL) and right (WR) halves of a birch leaf for calculation of fluctuating asymmetry are usually conducted perpendicular to midrib, in the middle of leaf lamina.

2.2. Experimental setup

The participants were haphazardly assigned to four groups. The control group received true information on the leaf origin (that all leaves were collected from the same tree); the other three groups received false information (Table 1). Therefore, although all participants in all four groups measured the same 100 leaves, the participants in each group had different beliefs regarding the sample origins, and therefore different expectations about the levels of FA that would likely be found in the samples.

We tested prediction 1 by haphazardly labelling 5 of 10 samples as collected from a polluted site and the other 5 samples as collected from an unpolluted site. These samples were offered to participants in group 2, and samples with interchanged labels were offered to participants in group 3 (Table 1). The recording of higher values of FA in the samples that were labelled as collected from a polluted site would indicate the existence of confirmation bias.

We tested prediction 2 by informing participants from group 4 that samples (labelled 1–10) were collected from both polluted and unpolluted sites (half and half), and then asking them to attribute each sample to either of these sites based on their own measurements of FA (Table 1). Samples measured by the control group were divided into two classes according to the decisions on sample origin made by the participants from group 4. A higher difference in FA between the two classes of samples reported by the participants from the experimental group relative to the measurements from the control group would indicate an influence of motivation of the observer to find differences in FA between polluted and unpolluted sites in the results of measurements.

2.3. Data analysis

We first conducted a mixed-model ANOVA (with 31 measurements of each of 100 leaves) to find out whether there was any evidence for directional asymmetry, size variation, and FA relative to measurement error (Palmer and Strobeck, 1986). Four of the 3100 FA values that exceeded 0.35 were excluded from the analyses, because they obviously resulted from an error in data entry. Individual values of FA reported by participants were first averaged within each sample, and then sample-specific means were averaged within two classes of samples ('polluted' and 'unpolluted').