

Precision and accuracy in quantifying herbivory

MARC T. J. JOHNSON, JEFFERY A. BERTRAND* and MARTIN M. TURCOTTE† Department of Biology, University of Toronto Mississauga, Mississauga, Canada

Abstract. 1. Tissue removal by herbivores (i.e. herbivory) is a dominant interaction in most communities which has important impacts on natural and managed ecosystems. Despite the importance of herbivory, we lack a quantitative comparison of the efficacy of the most commonly used methods used to quantify herbivore damage.

2. We examined the factors that affect the precision and accuracy of visual and digital methods commonly used to quantify damage to leaves.

3. We created 224 digital leaves from four plant species. In a fully factorial design we manipulated leaf morphology and species, the location of damage (marginal or internal), estimation method (exact percentage or 25% bins), observer experience and expectancy bias (i.e., bias due to an expected result). Using 583 adult observers, we estimated the precision and accuracy of individuals' ability to visually estimate known levels of damage. In a third smaller experiment, we performed similar analyses using a digital scanner.

4. Across the first two experiments, individuals estimated damage with high precision ($R^2 = 0.75$ and 0.80) and accuracy (slope_{actual vs estimated} = 0.88 and 0.86). However, the precision and accuracy of estimates were influenced by plant species, the location of damage, and estimation method. Inexperienced individuals also overestimated low levels of damage, and this bias decreased with experience. Digital methods were precise ($R^2 = 0.98$) whereas accuracy was statistically indistinguishable from visual methods (slope = 0.91).

5. Visual estimates of damage provide the fastest and most cost-effective method for quantifying herbivory, and our results show they can be precise and accurate. We use our results to provide specific recommendations for future research.

Key words. Crop damage, florivory, folivory, herbivore damage, methods in ecology, pest insect, plant-insect, primary consumption.

Introduction

The consumption of plants by arthropods and mammals (i.e. herbivory) is among the most important species interactions in nature. Animals consume a large proportion of plant primary production, and this herbivory represents the major conduit of energy into animal foodwebs (McNaughton *et al.*, 1989; Cyr & Pace, 1993; Turcotte *et al.*, 2014a,b). Herbivores can threaten the persistence of rare plants when they limit recruitment (Schemske

et al., 1994). They also help to explain the spread of exotic plants when introduced plants lose their herbivores in their new environment (Keane & Crawley, 2002; Agrawal & Kotanen, 2003; Colautti *et al.*, 2004), or when native or exotic herbivores selectively feed on native plants (Parker *et al.*, 2006; Kalisz *et al.*, 2014). In agriculture and forestry, herbivores can have large negative impacts on yield and regional economies (Ayres & Lombardero, 2000; Losey & Vaughan, 2006). Understanding the importance of herbivory in basic and applied ecology requires reliable estimates of herbivore damage. Unfortunately, standardised protocols to quantify herbivory do not exist because the accuracy and precision of commonly employed methods are rarely studied and thus poorly characterised (Lowman, 1984; Waller & Jones, 1989; Williams & Abbott, 1991).

Many modern studies of damage caused by herbivores and pathogens utilise digital analysis methods, including leaf area meters or scanners and image software (Lowman, 1984; Moles

Correspondence: Marc T. J. Johnson, Department of Biology, University of Toronto-Mississauga, 3359 Mississauga Rd., Mississauga, Ontario L5L 1C6, Canada. E-mail: marc.johnson@utoronto.ca

*Current address: 41 Warrender Avenue, Toronto, Ontario M9B 5Z2, Canada.

†Current address: Center for Adaptation to a Changing Environment, Institute of Integrative Biology, ETH Zürich, Universitätsstrasse 16, Zürich 8092, Switzerland.

& Westoby, 2000; Škaloudová *et al.*, 2006; Goodwin & Hsiang, 2010). Portable leaf area meters are commercially available from several companies (O'Neal *et al.*, 2002; Gong *et al.*, 2013) (e.g. LI-COR 3000C, Lincoln NE, USA; CI-202, CID Bio-Science, Camas, WA, USA; and AM350, Hoskin Scientific, Burlington, ON, Canada), and they are suitable for estimating changes in leaf area owing to growth or damage over time. Some instruments can also estimate the amount of herbivore or pathogen damage within leaf margins (i.e. internal damage). Unfortunately, all leaf area meters ignore damage along leaf margins, and thus they underestimate the amount of damage (O'Neal *et al.*, 2002). A more commonly employed digital method of estimating herbivore damage involves scanning leaves using a flatbed scanner and analysing the percentage of tissue damage using ImageJ, Adobe Photoshop, Scion or other software (Moles & Westoby, 2000; O'Neal *et al.*, 2002; Škaloudová *et al.*, 2006). Damage to the leaf margins still presents difficulties because researchers must retrace the leaf boundary based on reference leaves to determine how much of the leaf area has been removed (O'Neal *et al.*, 2002). A further limitation of scanners is the need to excise leaves, which prevents the study of plants *in situ* and may introduce undesirable confounds when leaves are subsequently given to insects in feeding trials (e.g. when leaves have glandular trichomes that are damaged during scanning). These difficulties combined with the costs of commercial leaf area meters, or the large handling time associated with scanning and analysing leaves (c. 40–70 s per leaf + collecting time) (O'Neal *et al.*, 2002), has limited the use of digital methods to estimate herbivory.

Visual estimates of leaves using either transparent grid paper (Coley, 1983; Lamarre *et al.*, 2012) or simple visual estimates (Stotz *et al.*, 2000; Johnson *et al.*, 2009a; Prasad *et al.*, 2012; Woods *et al.*, 2012; Turcotte *et al.*, 2014b; Lehdal & Ågren, 2015) remain the most commonly employed methods to estimate damage. The use of grid paper is expected to be precise and accurate but it is still time-consuming (Turcotte *et al.*, 2014b). In contrast, visual estimates (without the aid of grid paper) can be performed in the lab or on live plants in the field, they require no specialised equipment, and large sample sizes can be achieved with minimal time investment (10–20 s per leaf, including processing time). Damage is typically estimated by eye as the percentage of surface area removed, in which observers either estimate the amount of damage according to semi-quantitative bins (e.g. 0–25%, 26–50%, etc.) (Stotz *et al.*, 2000), or less often by estimating the exact percentage of tissue removed (Johnson *et al.*, 2009b). A problem with these approaches is that although convenient, the precision (i.e. repeatability of measurements) and accuracy (i.e. bias in estimating the true level of herbivory) of visual methods is questionable and rarely validated.

The objective of our study was to assess the precision and accuracy of visual and digital estimates of herbivore damage on leaves. We did this by creating digital model leaves (Fig. 1) with known amounts of simulated leaf herbivore damage (Fig. 2). Observers then visually estimated the percentage of tissue missing from individual leaves that varied in the amount of damage. We then used these data to quantify the precision and accuracy of each observer at estimating known levels of leaf area loss.

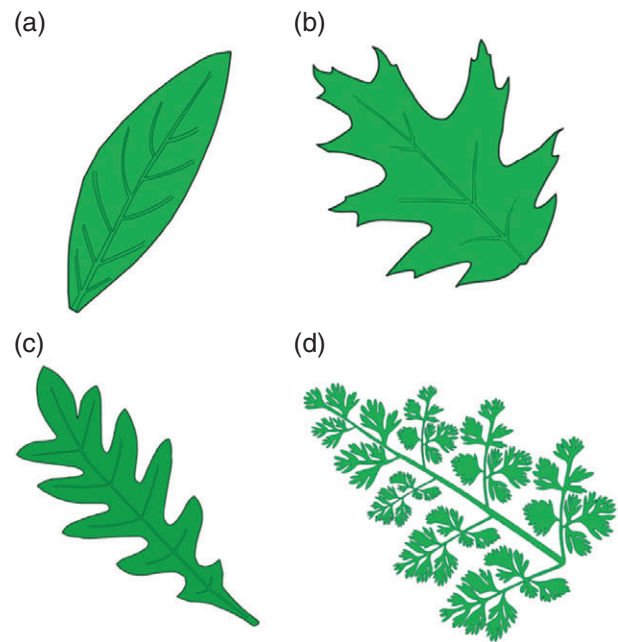


Fig. 1. Digital leaves of species used in the study. Depicted leaves represent: (a) common evening primrose (*Oenothera biennis*); (b) red oak (*Quercus rubra*); (c) shepherd's-purse (*Capsella bursa-pastoris*); and (d) queen anne's lace or wild carrot (*Daucus carota*). All leaves were created from scanned images of actual leaves that were outlined digitally and then rendered to 'forest green' and equal area. Experiment 1 used all species and Experiment 2 used only evening primrose and oak.

Using this approach we conducted three experiments. In Experiment 1, we assessed how the precision and accuracy of herbivory estimates of 19 scientists were influenced by: plant species that vary in leaf complexity (simple, lobed, dissected), damage type (internal vs marginal), estimation method (exact % vs. 25% bins), and the years of experience the observer had in measuring herbivory. In Experiment 2, we conducted a larger survey of 564 individuals to assess how leaf complexity, damage type, method of estimation, and an individuals' expectancy bias (i.e. bias due to an expected result) affected the precision and accuracy of herbivory estimates. In Experiment 3, we assessed the accuracy and precision of digital measurements of herbivory. We use our results to make specific recommendations on how to maximise the precision and accuracy of quantifying damage to leaves.

Materials and methods

Study design and digital leaves

To determine the precision and accuracy of visual estimates, we conducted surveys of individuals affiliated with the lead author's lab (Experiment 1), plus students enrolled in two undergraduate classes at the University of Toronto at Mississauga (Experiment 2). During these surveys, participants were presented with digital representations of damaged leaves from four North American plant species: *Oenothera biennis* L. (common evening primrose, Onagraceae, Fig. 1a), *Quercus rubra*

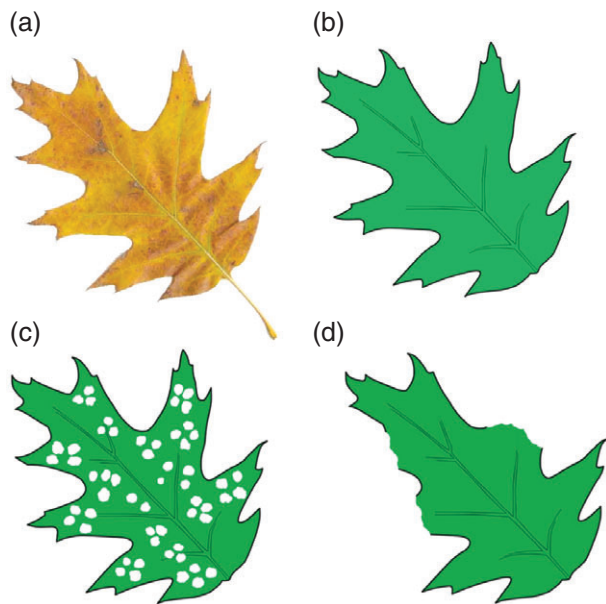


Fig. 2. Examples of how digital oak leaves were created and manipulated to simulate damage. (a) We first scanned an undamaged representative leaf of a given species. (b) We then outlined the leaf and converted it to forest green. (c) Internal damage was made by removing small, irregular-shaped holes from the interior of the leaf. This type of damage is consistent with the damage made by many small caterpillars and some beetles (e.g. Alticinae, Curculionidae). (d) 'Marginal damage' was made by removing portions of the leaf margin; this pattern of damage is consistent with the type of damage caterpillars and some large beetles might make.

L. (red oak, Fagaceae, Fig. 1b), *Capsella bursa-pastoris* L. shepherd's-purse, Brassicaceae, Fig. 1c), and *Daucus carota* (L.) Medid. (wild carrot, Apiaceae, Fig. 1d). These species were chosen as models because they vary in leaf shape complexity, and they are common in the region of the study. Complexity reflected the presence and number of leaf lobes, as well as the presence or absence of leaflets (Fig. 1).

Digital representations were chosen over real leaves because they allowed us to control for individual differences in leaf texture, colour, size, shape, as well as the location and type of damage (Fig. 2). Although some realism is lost in comparison to actual leaves, we could not have achieved the replication, robust experimental design, and precise manipulations of damage without digitisation. Digital leaves were constructed using Adobe Photoshop CS5 Extended Edition (Adobe Systems, San Jose, California). Leaves were constructed by first scanning a typical leaf of each species (Fig. 2a) and digitally tracing the leaf to create a new digital image that was rendered in a solid 'forest green' colour (Fig. 2b). We adjusted leaves to a uniform surface area.

A total of 224 digital leaves were constructed, 56 per species, subdivided into 28 leaves per species with only internal damage (Fig. 2c) and 28 per species with only marginal damage (Fig. 2d). Damage levels ranged from 1% to 97%, where damage was created by removing portions of the leaf area in a way that mimicked the appearance of actual damage. Internal damage was restricted to the interior of the leaf margin, and it resembled

chewing damage by small caterpillars, flea beetles, and weevils (Fig. 2c). Marginal damage was restricted to leaf margins, and it resembled chewing damage by many caterpillars and some beetle species (Fig. 2d).

Experiment 1

In Experiment 1, 19 individuals estimated the amount of damage on all 224 digital leaves from the four species. The individuals surveyed included undergraduate students, graduate students, post-doctoral fellows, and one faculty member (M. Johnson) affiliated with the lead PI's lab. The virtual leaves were created by J. Bertrand, who administered all surveys. Individuals ranged from having over 10 years of experience studying plant–herbivore interactions and estimating herbivory, to no prior research experience. Each person was presented with either a 28-page hard copy survey featuring eight damaged leaves per page from evening primrose, red oak, shepherd's-purse, or carrot, or the same leaves were displayed as a PowerPoint presentation on a large TV monitor with four leaves per slide. Participants were provided with written instructions as well as a page containing undamaged reference leaves (Appendix S1). Leaves were presented in random order and individuals were given 10–15 s to estimate the damage to each leaf. This method was used to simulate the typical time constraints one faces when collecting large amounts of data in the field or lab from multiple plant species.

We designed the survey as a fully factorial experiment that manipulated leaf species (i.e. evening primrose, oak, shepherd's-purse, and carrot), damage type (marginal or internal), and estimation method (exact % vs. bins; see below). We also used the number of years of experience as a covariate in analyses. To manipulate the method of estimation, we instructed individuals to either estimate damage as an exact percentage or within one of four bins (0–25%, 26–50%, 51–75%, and 76–100%).

Experiment 2

To build on our results from Experiment 1, we conducted a second experiment with more observers, but only two of the leaf species used previously (evening primrose and oak). We surveyed students in each of two undergraduate biology classes at the University of Toronto at Mississauga; a first-year biology class (429 students) and a third-year evolutionary biology class (135 students). A total of 60 leaves were presented to students via a PowerPoint presentation projected on a large screen, which included 30 leaves of each species equally split between marginal and internal damage. Unlike the previous experiment, we first presented all 30 leaves of evening primrose to the class, followed by the 30 oak leaves. Evening primrose and red oak were used because they represented variation in leaf complexity where the former species was the simplest leaf from Experiment 1, and the oak leaf was the leaf species that individuals had the greatest difficulty in estimating during our first experiment. Leaves were presented in a randomised order with respect to the amount and type of damage. Each slide featured four damaged leaves of either evening primrose or oak along with

an undamaged leaf from a given species as a reference; leaf area was kept constant minus the damage. Students from the third year class were given 60 s to examine leaves on each slide (15 s per leaf), and they were asked to record the percentage of tissue removed from each leaf. Students from the first year class were given only 40 s per slide (10 s per leaf) because the initial survey showed that most students finished their estimates within 30 s.

Experiment 2 was also designed as a fully factorial experiment in which we manipulated leaf species (evening primrose or oak), method of estimation (exact percentage or bins), damage type (marginal or internal), and expectancy bias (see below). As with Experiment 1, an equal number of students were asked to estimate damage as either an exact percentage or according to four percentage bins (see above). We also sought to understand whether a student's estimation of damage was biased by an *a priori* hypothesis. To test for this bias, students were given instructions in writing stating that the digital leaves were rendered from actual leaves with real damage. One-third of students were instructed that they were testing the hypothesis that lobed leaves experience greater damage from herbivores than unlobed leaves (oaks > evening primrose, Appendix S2). A second one-third of students were informed that they were testing the hypothesis that unlobed leaves experience greater damage from herbivores than lobed leaves (evening primrose > oaks). The remaining students were not informed of any *a priori* hypothesis. The class (first- or third-year biology) was treated as a blocking factor in analyses.

Experiment 3

We assessed the precision and accuracy of the most common digital scanning method. Researchers frequently scan leaves on a flatbed scanner and use various image analysis software to estimate the amount of tissue removed by herbivores (O'Neal *et al.*, 2002). This can be accomplished easily and accurately when there is only internal damage that does not border a leaf margin because it is easy to subtract the internal area removed from the total area bound by leaf margins. Difficulties arise when herbivores remove leaf margins, which is often the case. In these instances, researchers must attempt to retrace the original leaf margin using either reference leaves or their intuitive sense of where the margin would have originally been. To assess the precision and accuracy of this latter digital method, we conducted an experiment in which the lead author (M. Johnson) retraced the margins of 30 *O. biennis* and 30 *Q. rubra* leaves featuring various degrees of marginal damage made by J. Bertrand in a way that was blind to the lead author. Thirty leaves of each species were presented to M. Johnson in a random order. All of these leaves had marginal damage ranging from 5% to 94%, and these leaves were multiplied by a randomly determined size factor between 40% and 175% to emulate natural variation in leaf size. No variation in leaf shape was made, and so this test probably represents the 'best case scenario' – an anti-conservative estimate of the precision and accuracy of digital methods. After tracing, leaf images were scanned using an HP Officejet 7500 desktop scanner and analysed using the Magic Wand and Analysis tools present in Adobe Photoshop CS5 Extended Edition to estimate the leaf area removed.

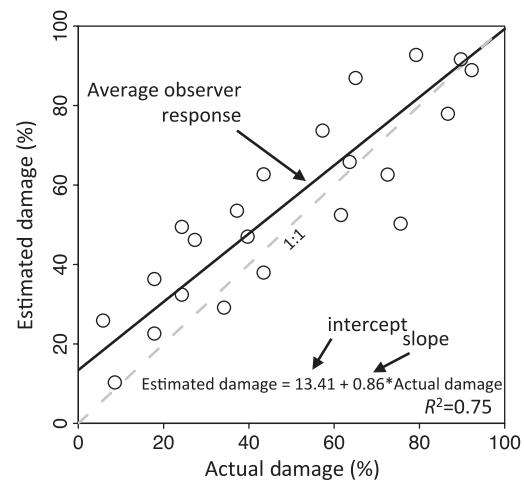


Fig. 3. Example of how precision (R^2) and accuracy (slope and intercept) were estimated by regressing an observer's 'estimated damage' against the 'actual damage' of multiple leaves. For each observer within each treatment group and plant species, we regressed the percentage herbivore damage they estimated against the actual % damage we removed from digital leaves. We then fit these data to a linear equation (black solid line) and estimated the R^2 , slope, and intercept. If an observer perfectly estimated damage than their observations would fall along the 1 : 1 line (dashed grey line) with slope = 1, $R^2 = 1$ and intercept = 0. The black line, equation and R^2 show the overall average slope, intercept and R^2 from Experiment 2. The data points shown are fictitious and included only to give readers a sense of how the data was generated and estimated.

Statistical analyses

We estimated the precision and accuracy of each observer by regressing the estimated damage against the actual damage across replicate leaves. This was done separately for each treatment combination of the damage type \times plant species interaction. Thus, for Experiment 1 there were eight separate regressions (i.e. two damage types \times four species) for each observer, and each regression consisted of 28 estimates of damage. For Experiment 2, there were four regressions per individual (i.e. two damage types \times two species) that consisted of 15 data points per regression. Regressions involving bins utilised the midpoint value of each bin. From each regression, we quantified precision as the coefficient of determination (R^2) (Fig. 3), which estimates the proportion of the variation in actual damage that was accurately captured by visual estimates. We quantified accuracy or bias as the slope of the line (β), where $\beta = 1$ indicates no bias, $\beta \neq 1$ represents over or under-estimation of the actual amount of damage. We also estimated the intercept (β_0), which quantifies the bias at low levels of actual damage (Fig. 3). We estimated 152 and 2250 estimates (one estimate for each level of a treatment combination per individual) of each response variable (R^2 and β) in Experiments 1 and 2, respectively.

Using the two response variables, we performed mixed model analysis of variance (ANOVA) using the *lme4* (Bates *et al.*, 2014) and *lmerTest* (Kuznetsova *et al.*, 2014) packages in R version 3.02 (R Core Team, 2013). *lmerTest* uses the algorithms of SAS Proc Mixed (Littell *et al.*, 1996) to determine the significance of fixed effects using type 3 sums-of-squares.

We also used this package to adjust the denominator d.f. using the Kenward–Roger correction method (Kenward & Roger, 1997). In these models, all effects were treated as fixed effects except student, which was a random effect nested within either Estimation Method (Experiment 1) or the Estimation Method \times Expectancy Bias interaction (Experiment 2). The small number of students in Experiment 1 made it impossible to run the full model with all main effects and interactions in the model. We, therefore, started with all main effect terms in the model and added two-way interactions using a forward stepwise selection procedure. Specifically, we tested the significance of each term individually, and we included the interaction with the lowest P -value < 0.1 . We then reran this selection process on the remaining interaction terms conditional on the inclusion of the first significant interaction term.

Experiment 2 had much larger sample sizes, so we used a backward model selection procedure, where we started with all main effects and two-way interactions included in the model. Terms were removed sequentially based on having the lowest statistical significance > 0.1 . We retained all main effects in the final model regardless of the P -value. We also included any significant interactions with class, even though we were not inherently interested in this interaction.

In Experiment 3, we conducted an analysis of variance with type-III SS significance tests implemented using the *Anova* function from the *car* package (Fox & Weisberg, 2011). In this analysis, we regressed the estimated amount of damage against the actual amount of damage, while testing whether there was an effect of plant species (evening primrose vs oak) or a damage \times species interaction.

Before starting analyses, we examined whether the starting models met assumptions of homogeneity of variance and normality and applied transformations to the data to remedy any deviations from these assumptions. For Experiment 1, we transformed R^2 and β with x^2 and $x^{0.5}$ transformations, respectively. For Experiment 2, we used an x^2 transformation for both variables. No transformation was used for Experiment 3.

Results

Experiment 1: accuracy and precision of estimating damage visually from species with varying leaf complexity

In this experiment, we assessed how prior research experience, leaf species, method of estimation, and damage type affected the precision and accuracy of 4256 visual estimates of herbivory by 19 individuals. Overall, individuals estimated leaf herbivory with a high precision (mean $R^2 = 0.802$) whereas accuracy measured as the slope of the line was significantly less than 1 ($\beta_{\text{mean}} = 0.880$; 95% CI = 0.859, 0.900). The intercept of the line averaged across individuals was significantly greater than 0 ($\beta_0 = 9.16\%$; 95% CI = 6.67, 11.65), indicating that individuals typically overestimated damage at low levels, but this overestimation decreased at higher damage because $\beta_{\text{mean}} < 1$ (see example in Fig. 3). We detected no effect of leaf species or an individual's prior experience in estimating herbivory on precision ($P > 0.2$, Table 1a). There was a significant interaction

Table 1. Analysis of variance of factors affecting the (a) precision (R^2) and (b) accuracy (slope) of visually estimating herbivory in Experiment 1.

Effect	d.f.	SS	MS	F	P
(a) R^2					
Experience	1,16	0.015	0.015	0.064	0.803
Species	3,128	0.094	0.031	1.406	0.244
Method	1,16	0.040	0.040	1.571	0.228
Damage type	1,128	0.068	0.068	0.003	0.959
Method \times dam. type	1,128	0.125	0.125	5.630	0.019
(b) Slope					
Experience	1,16	0.010	0.010	0.670	0.425
Species	3,126	0.079	0.026	4.828	0.003
Method	1,16	0.018	0.018	5.363	0.034
Damage type	1,126	0.041	0.041	13.868	<0.001
Species \times method	3,126	0.024	0.008	2.697	0.049

For each response variable we tested the effects of experience of the observer, leaf species (*Oenothera biennis*, *Quercus rubra*, *Capsella bursa-pastoris*, *Daucus carota*), method of estimation (exact percentage or bins), damage type (marginal versus internal), and significant interactions that entered into the model during forward stepwise selection. For each effect we show the sums-of-squares (SS), mean-square (MS), degrees-of-freedom (d.f.), F - and P -value. Denominator d.f. were adjusted according to the Kenward–Roger correction method for small sample sizes. Significant effects at $P < 0.05$ are shown in bold.

between the method of estimation (exact vs. 25% bins) and damage type (marginal vs. internal) for precision (Table 1a; Fig. 4a). This interaction occurred because R^2 values were 12% higher for exact estimates versus % bins for marginal damage ($P = 0.03$), whereas there was no difference in precision for internal damage.

The accuracy of herbivory estimates (i.e. slope or β) was influenced by damage type, leaf species, and estimation method (Table 1B). Marginal damage was estimated with 7% greater accuracy than internal damage ($\beta_{\text{marginal}} = 0.888$, $\beta_{\text{internal}} = 0.825$) and exact estimates of damage ($\beta = 0.900$) had a 10% higher slope than binned estimates ($\beta = 0.817$, Fig. 4b). The accuracy of herbivory estimates varied by as much as 12% among leaf species, but there was no clear relationship with leaf complexity; herbivory on oaks was estimated with the lowest accuracy ($\beta = 0.795$), whereas damage on evening primrose ($\beta = 0.871$), shepherd's-purse ($\beta = 0.892$), and wild carrot (0.862) were estimated with similar levels of accuracy. An interaction between the estimation method and species was caused by the difference between exact and binned estimation methods being greater in carrot and evening primrose than in shepherd's-purse and oak (Fig. 4b). Although there was no effect of experience on the slope of the line, the intercept (i.e. bias at low actual herbivory) was significantly negatively correlated ($r = -0.74$, $P < 0.001$; Fig. 5), such that individuals with 10 or more years of experience estimated low levels of herbivory with little or no bias.

Experiment 2: accuracy and precision of visually estimating damage from two species

After the results from Experiment 1, we conducted a larger experiment (564 individual observers; 33 685 herbivory

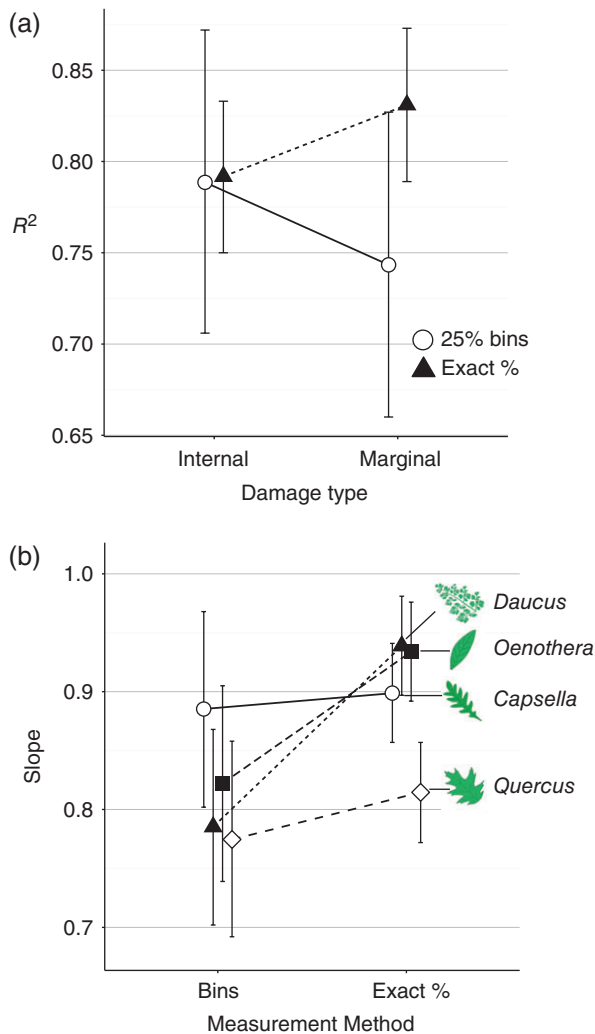


Fig. 4. Effects of damage type, estimation method, and species on observers' precision (R^2) and accuracy (slope) when estimating damage in Experiment 1. (a) Damage type (internal vs. marginal) and estimation method (exact percentage vs. 25% bins) interacted ($P=0.019$) to affect R^2 . (b) The estimation method also interacted with plant species ($P=0.049$) to affect the slope of the relationship between estimated and actual damage.

estimates) to better understand how leaf species, estimation method, damage type as well as an observer's *a priori* bias affect herbivory estimates. Given that researcher experience had no effect on the precision or accuracy (slope) in Experiment 1, we tested the ability of students with little to no research experience from two large undergraduate classes to estimate herbivory from two species with contrasting leaf shapes (evening primrose – simple leaves; red oak, lobed and toothed leaves; Fig. 1).

Multiple factors affected the precision of herbivory estimates. Overall, the precision of estimates (mean $R^2=0.751$; Fig. 6a) was lower in Experiment 2 than Experiment 1. Damage type had the largest effect on precision with R^2 being 13% higher for estimates of marginal damage ($R^2=0.810$) than internal damage

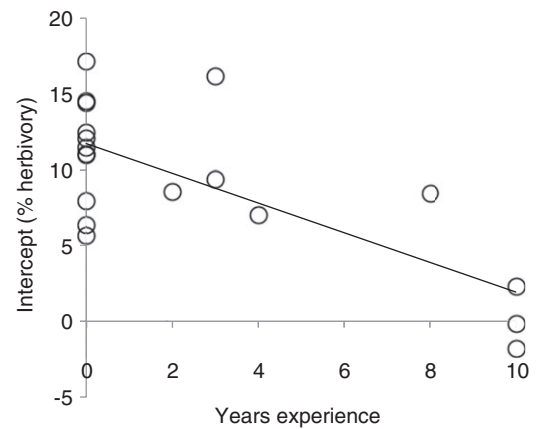


Fig. 5. Low levels of herbivory are overestimated by individuals with less experience. The intercept was estimated from the equation of the line of estimated % damage versus actual damage. The intercept represents the % damage estimated by individuals when the actual level of damage is 0. The intercept was negatively correlated with experience of the observer ($r=-0.74$, $P<0.001$).

($R^2=0.714$, Table 2a). This effect was even stronger for evening primrose than oak, which resulted in a significant damage type \times species interaction (Fig. 6b; Table 2a). Estimating damage as an exact percentage was associated with 4.3% higher R^2 values than using 25% bins ($R^2_{\text{exact \%}}=0.778$, $R^2_{\text{bins}}=0.746$ Table 2a, see Estimation Method). Biasing students' expectations about whether lobed or unlobed leaves are expected to have more herbivory did not impact precision. Finally, the two classes differed in their precision, with the third year undergraduate class having 6% higher R^2 values than the first-year class. The class also interacted with leaf species and damage type, but these interactions tended only to magnify the magnitude of the main effect in the third-year class versus the first-year class.

As with Experiment 1, multiple factors influenced the accuracy of herbivory estimates. The overall accuracy ($\beta=0.861$; 95% CI=0.851, 0.871; Fig. 7a) was similar to Experiment 1, whereas the intercept was even greater ($\beta_0=14.14\%$; 95% CI=13.31, 14.96), showing again that individuals overestimated low levels of damage and this bias decreased at higher levels of damage. The estimation method had the largest effect on precision, with exact estimates yielding 4% higher slopes than binned estimates. As in Experiment 1, this effect of the estimation method was greater for internal damage than marginal damage (Fig. 7c). Damage type and leaf species both had small yet significant effects on accuracy, with the slope being 1% higher for internal than marginal damage, and 1% higher for evening primrose than oak. Damage type and species also interacted with internal damage having 1.4% higher slopes than marginal damage on evening primrose (Fig. 7b), whereas marginal damage had 4% higher slopes than internal damage on oak (Fig. 7b). The effect of the class also had a significant effect on accuracy, with the third year class having 4% higher slopes (0.904) than the first year class (0.869). This latter effect may reflect the greater amount of time given to the third-year class to assess damage on each leaf.

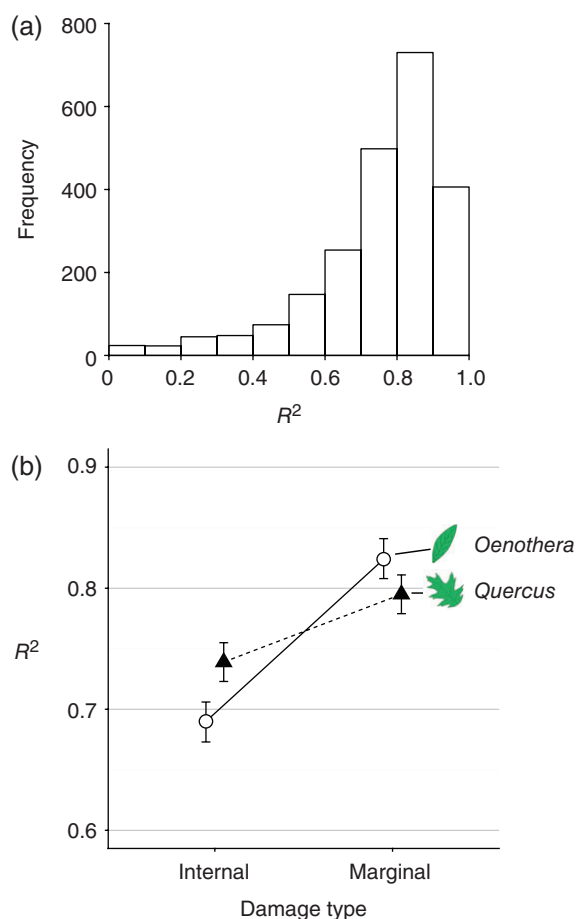


Fig. 6. Variation in precision (R^2) and the effects of damage type and species on R^2 in Experiment 2. (a) Observers estimated damage with fairly high precision. (b) Damage type and species interacted to affect accuracy when estimating damage ($P < 0.001$).

Experiment 3: accuracy and precision of estimating marginal damage by scanning

In Experiment 3, we estimated the bias and accuracy of scanning leaves to estimate damage. When leaves are scanned, observers must first reconstruct the damaged leaf margin before estimating the percentage of tissue missing. We did this for 60 digitally created leaves of varying sizes, split equally between evening primrose and oak (see Materials and methods). There was no difference between species in the ability to estimate herbivory ($F_{1,37} = 2.32$, $P = 0.13$). The traced and scanned leaves predicted the variation in the actual damage with very high precision ($R^2 = 0.978$), whereas estimates were biased by 9% ($\beta = 0.911$, 95% CI: 0.859–0.963). Low levels of damage were slightly over-estimated ($\beta_0 = 2.548\%$), but this estimate did not significantly deviate from 0 ($t_{56} = 1.696$, $P = 0.095$).

Discussion

The present study shows that visual estimates of herbivore damage can be precise and accurate, but aspects of a system's

Table 2. Analysis of variance of factors affecting the (a) precision (R^2) and (b) accuracy (slope) of visual estimates of herbivory in Experiment 2.

Effect	d.f.	SS	<i>F</i>	<i>P</i>
(a) R^2				
Class	1,557	0.617	19.246	<0.001
Bias	2,559	0.029	0.446	0.641
Species	1,1681	0.205	0.006	0.940
Method	1,559	0.538	17.321	<0.001
Damage type	1,1681	14.238	266.266	<0.001
Class \times species	1,1681	0.513	16.485	<0.001
Class \times dam. type	1,1681	0.443	14.237	<0.001
Species \times dam. type	1,1681	1.430	45.975	<0.001
(b) Slope				
Class	1,558	0.544	13.195	<0.001
Bias	2,559	0.037	0.459	0.632
Species	1,1680	0.000	5.636	0.018
Method	2,559	0.889	22.211	<0.001
Damage type	1,1680	0.004	6.901	0.009
Class \times species	1,1680	0.847	21.267	<0.001
Class \times dam. type	1,1680	1.263	30.910	<0.001
Species \times dam. type	1,1680	0.639	15.999	<0.001
Method \times dam. type	1,1680	1.193	29.979	<0.001

For each response variable, we tested the effects of class (third year versus first year biology), *a priori* bias, leaf species (*Oenothera biennis*, *Quercus rubra*), estimating method ('Method' - exact percentage or bins), damage type (marginal versus internal), and significant interactions that remained in the model after backward stepwise selection. For each effect we show the sums-of-squares (SS), mean-square (MS), degrees-of-freedom (d.f.), *F*- and *P*-value. Denominator degrees of freedom were adjusted according to the Kenward–Roger correction method. Significant effects at $P < 0.05$ are shown in bold.

biology and method of estimation can reduce the utility of these methods. Several results from our study are most important to our conclusions. First, leaf morphology and the type of damage imposed by herbivores (i.e. internal versus marginal damage) influence both the precision and bias of damage. Second, exact estimates of damage consistently lead to greater precision and accuracy than using coarser-binned estimates of damage. Third, individuals with little or no experience can estimate damage precisely, but there was large variation among observers and individuals with less experience consistently overestimate low levels of damage. Finally, digital methods provide the most precise estimates of damage but their level of bias is statistically equivalent to visual estimates of damage. We discuss the implications of these results below and provide practical recommendations for measuring herbivore damage.

Precision and accuracy of visual estimates of herbivory

Particular aspects of the biology of both plants and their herbivores influence the precision and accuracy of visual estimates of herbivory. A plant's leaf morphology has a significant impact on estimates of herbivory. We found no clear relationship between leaf complexity and our metrics of precision (R^2) or accuracy (slope) (Fig. 3), but leaves with lobes and teeth (i.e. oak) were harder to accurately measure (Figs 4b, 6b and 7b), especially

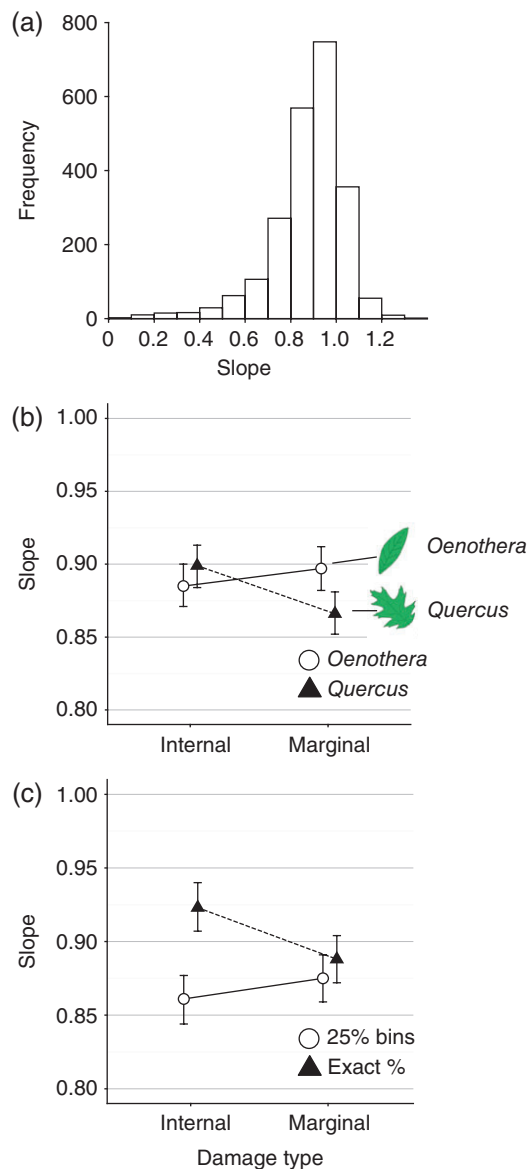


Fig. 7. The effects of damage type, species, and estimation method on observer accuracy (slope) in Experiment 2. (a) Observers estimated damage with high accuracy. Damage type interacted with: (b) species ($P < 0.001$) and (c) estimation method to affect accuracy ($P < 0.001$).

when damage occurred along leaf margins and exact estimates of damage were used. The combination of lobes and teeth likely make it difficult for the eye to distinguish damaged from intact leaf margins. This effect may be even greater if one accounts for variation in leaf shape within species, which was fixed in our study for simplicity. The location of herbivory (along margins or internal) also affects precision and accuracy. The precision of estimates was highest when the damage occurred along margins, and this effect was most pronounced when exact measures of damage were used. The location of damage did influence the accuracy of estimates, but these effects depended on the plant species' leaf morphology, and they varied between experiments.

The methods used to estimate herbivory also have a large influence on the precision and accuracy of visual estimates of herbivory. Estimating herbivory as an exact percentage provides the most precise and least biased method for estimating damage. Many studies estimate damage on a coarser scale by either counting the proportion of leaves that are damaged (Johnson *et al.*, 2009a; Woods *et al.*, 2012), or categorising damage on individual leaves into broad bins of per cent damage or ordinal scores of the level of damage (Pilson, 2000; Van Zandt & Agrawal, 2004). Our results show that such practices should be avoided, especially given that most plant species experience $< 10\%$ leaf damage annually (Turcotte *et al.*, 2014a). Thus, many studies use bins that are too large to distinguish accurately rates of damage among groups. Our results in Experiment 2 also suggest that the best results are obtained when adequate time is provided to scan leaves visually (Table 2; see Effect of Class). Finally, although digital scanners can be very precise, their estimates are similarly biased when damaged leaf margins need to be reconstructed.

Our findings build on past studies that have sought to identify the best methods for estimating the damage caused by herbivores. It has long been appreciated that single estimates of herbivore damage (i.e. snapshot or standing measures) can underestimate damage especially when herbivores consume entire leaves or when damaged leaves drop prematurely (Lowman, 1984; Coley & Barone, 1996). In such cases, repeated measurements of marked leaves provide the most accurate estimate of the percentage of tissue removed. Changes in the timing and rate of herbivory, and the longevity and turnover of leaves can also influence estimates of herbivory, which has led to methods that explicitly incorporate these parameters (Sand-Jensen *et al.*, 1994; Lamarre *et al.*, 2012). Finally the size, thickness, and density of leaves frequently vary within and between species and failing to account for such variation can lead to biased estimates of damage (Waller & Jones, 1989; Williams & Abbott, 1991). Our study complements these past methodological developments. Most importantly, our results show that the use of expensive digital area meters or time-consuming flatbed scanners is unnecessary in many applications.

Recommendations

Our results lead to several recommendations about how and when to implement visual estimates of the amount of damage to plants by herbivores:

1. Visual estimates can be precise and accurate. Simply scanning a leaf by eye provides the fastest and most cost-effective method for estimating damage from large numbers of samples. Visual estimates will be most useful for comparative purposes, but if precise estimates of damage are required (e.g. when parameterising models of energy flux), then digital methods are often superior albeit not bias-free.
2. Training is critical. Although individuals with no prior experience can provide precise estimates of damage, there is large variation among individuals in their precision and

accuracy (see Figs 6a and 7a). Furthermore, inexperienced observers often overestimate low levels of damage (Fig. 5). We suggest providing individuals with clear instructions on how to estimate damage. Individuals should first examine a large number of leaves to gain a sense of the range of leaf shape, size, and damage. To quantify herbivory, individuals should mentally break a leaf into sections (e.g. halves, thirds, and quarters) and estimate damage to those sections separately. Individuals should then be given a training set of leaves to estimate and improve their precision and bias. To obtain the training sets used in our experiments, please contact the corresponding author.

3. Slower is better. We changed the amount of time given per leaf between the two large classes (15 vs. 10 s per leaf) and observed a significant decrease in both precision and accuracy. We recommend 15–20 s per leaf for the best results, which is still substantially faster than digital methods.
4. Average estimates across multiple leaves per plant. Precision and accuracy will be improved when individuals' estimate damage from multiple leaves (e.g. 10–20) per plant and average their estimates. The optimal number of leaves can be determined with pilot data that assesses the sample size required to obtain a mean value with minimum variance. This application makes the explicit assumption that leaves are generally of the same size (Williams & Abbott, 1991).

Conclusions

Our large-scale experiments show that visual estimates of herbivory provide a valid method to estimate damage accurately. Some biases do exist with both visual and digital methods, which are associated with the biology of the plant and herbivore species, as well as the methods employed. These biases can be minimised by following the recommendations presented in this paper.

Acknowledgements

We are indebted to Professor Richter and the BIO153 and BIO342 classes for assisting us with Experiment 2. The members of M. Johnson's EvoEco Lab also kindly made themselves available for the collection of data in Experiment 1. The manuscript was improved by comments from editor T. Ohgushi and two anonymous reviewers. This work was supported by an NSERC Discovery Grant and Connaught Early Researcher Award to M. Johnson.

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: 10.1111/een.12280

Appendix S1. Instructions provided to observers in Experiment 1.

Appendix S2. Instructions provided to observers in Experiment 2.

References

- Agrawal, A.A. & Kotanen, P.M. (2003) Herbivores and the success of exotic plants: a phylogenetically controlled experiment. *Ecology Letters*, **6**, 712–715.
- Ayres, M.P. & Lombardero, M.J. (2000) Assessing the consequences of global change for forest disturbance from herbivores and pathogens. *Science of the Total Environment*, **262**, 263–286.
- Bates, D., Maechler, M., Bolker, B. & Walker, S. (2014) *lme4: Linear mixed-effects models using Eigen and S4*, Version 1.0-6 [WWW document]. URL <http://CRAN.R-project.org/package=lme4> [accessed on 22 September 2014].
- Colautti, R.I., Ricciardi, A., Grigorovich, I.A. & MacIsaac, H.J. (2004) Is invasion success explained by the enemy release hypothesis? *Ecology Letters*, **7**, 721–733.
- Coley, P.D. (1983) Herbivory and defensive characteristics of tree species in a lowland tropical forest. *Ecological Monographs*, **53**, 209–233.
- Coley, P.D. & Barone, J.A. (1996) Herbivory and plant defenses in tropical forests. *Annual Review of Ecology and Systematics*, **27**, 305–335.
- Cyr, H. & Pace, M.L. (1993) Magnitude and patterns of herbivory in aquatic and terrestrial ecosystems. *Nature*, **361**, 148–150.
- Fox, J. & Weisberg, S. (2011) *An R Companion to Applied Regression*, 2nd edn. Sage, Thousand Oaks, California.
- Gong, A., Wu, X., Qiu, Z. & He, Y. (2013) A handheld device for leaf area measurement. *Computers and Electronics in Agriculture*, **98**, 74–80.
- Goodwin, P.H. & Hsiang, T. (2010) Quantification of fungal infection of leaves with digital images and Scion Image software. *Molecular and Cell Biology Methods for Fungi*, pp. 125–135. Springer, New York, New York.
- Johnson, M.T.J., Agrawal, A.A., Maron, J.L. & Salminen, J.-P. (2009a) Heritability, covariation and natural selection on 24 traits of common evening primrose (*Oenothera biennis*) from a field experiment. *Journal of Evolutionary Biology*, **22**, 1295–1307.
- Johnson, M.T.J., Smith, S.D. & Rausher, M.D. (2009b) Plant sex and the evolution of plant defenses against herbivores. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 18079–18084.
- Kalisz, S., Spigler, R.B. & Horvitz, C.C. (2014) In a long-term experimental demography study, excluding ungulates reversed invader's explosive population growth rate and restored natives. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 4501–4506.
- Keane, R.M. & Crawley, M.J. (2002) Exotic plant invasions and the enemy release hypothesis. *Trends in Ecology & Evolution*, **17**, 164–170.
- Kenward, M.G. & Roger, J.H. (1997) Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, **53**, 983–997.
- Kuznetsova, A., Brockhoff, B. & Christensen, H.B. (2014) *lmerTest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package)*, Version 2.0-11 [WWW document]. URL <http://CRAN.R-project.org/package=lmerTest> [accessed on 22 September 2014].
- Lamarre, G.P.A., Baraloto, C., Fortunel, C., Dávila, N., Mesones, I., Rios, J.G. *et al.* (2012) Herbivory, growth rates, and habitat specialization in tropical tree lineages: implications for Amazonian beta-diversity. *Ecology*, **93**, S195–S210.
- Lehndal, L. & Ågren, J. (2015) Latitudinal variation in resistance and tolerance to herbivory in the perennial herb *Lythrum salicaria* is related to intensity of herbivory and plant phenology. *Journal of Evolutionary Biology*, **18**, 576–589.

- Littell, R.C., Milliken, G.A., Stroup, W.W. & Wolfinger, R.D. (1996) *SAS System for Mixed Models*. SAS Institute, Cary, North Carolina.
- Losey, J.E. & Vaughan, M. (2006) The economic value of ecological services provided by insects. *Bioscience*, **56**, 311–323.
- Lowman, M.D. (1984) An assessment of techniques for measuring herbivory: is rainforest defoliation more intense than we thought? *Biotropica*, **16**, 264–268.
- McNaughton, S.J., Oesterheld, M., Frank, D.A. & Williams, K.J. (1989) Ecosystem-level patterns of primary productivity and herbivory in terrestrial habitats. *Nature*, **341**, 142–144.
- Moles, A.T. & Westoby, M. (2000) Do small leaves expand faster than large leaves, and do shorter expansion times reduce herbivore damage? *Oikos*, **90**, 517–524.
- O'Neal, M.E., Landis, D.A. & Isaacs, R. (2002) An inexpensive, accurate method for measuring leaf area and defoliation through digital image analysis. *Journal of Economic Entomology*, **95**, 1190–1194.
- Parker, J.D., Burkepile, D.E. & Hay, M.E. (2006) Opposing effects of native and exotic herbivores on plant invasions. *Science*, **311**, 1459–1461.
- Pilson, D. (2000) Herbivory and natural selection on flowering phenology in wild sunflower, *Helianthus annuus*. *Oecologia*, **122**, 72–82.
- Prasad, K.V.S.K., Song, B.-H., Olson-Manning, C., Anderson, J.T., Lee, C.-R., Schranz, M.E. *et al.* (2012) A gain-of-function polymorphism controlling complex traits and fitness in nature. *Science*, **337**, 1081–1084.
- Sand-Jensen, K., Jacobsen, D. & Duarte, C.M. (1994) Herbivory and resulting plant damage. *Oikos*, **69**, 545–549.
- Schemske, D.W., Husband, B.C., Ruckelshaus, M.H., Goodwillie, C., Parker, I.M. & Bishop, J.G. (1994) Evaluating approaches to the conservation of rare and endangered plants. *Ecology*, **75**, 584–606.
- Škaloudová, B., Křivan, V. & Zemek, R. (2006) Computer-assisted estimation of leaf damage caused by spider mites. *Computers and Electronics in Agriculture*, **53**, 81–91.
- Stotz, H.U., Pittendrigh, B.R., Kroymann, J.R., Weniger, K., Fritsche, J., Bauke, A. *et al.* (2000) Induced plant defense responses against chewing insects. Ethylene signaling reduces resistance of *Arabidopsis* against Egyptian cotton worm but not diamondback moth. *Plant Physiology*, **124**, 1007–1018.
- R Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing [WWW document]. URL: <http://www.R-project.org> [accessed on 22 September 2014].
- Turcotte, M.M., Davies, T.J., Thomsen, C.J.M. & Johnson, M.T.J. (2014a) Macroecological and macroevolutionary patterns of leaf herbivory across vascular plants. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20140555.
- Turcotte, M.M., Thomsen, C.J.M., Broadhead, G.T., Fine, P.V.A., Godfrey, R.M., Lamarre, G.P.A. *et al.* (2014b) Percentage leaf herbivory across vascular plant species. *Ecology*, **95**, 788–788.
- Van Zandt, P.A. & Agrawal, A.A. (2004) Specificity of induced plant responses to specialist herbivores of the common milkweed, *Asclepias syriaca*. *Oikos*, **104**, 401–409.
- Waller, D.A. & Jones, C.G. (1989) Measuring herbivory. *Ecological Entomology*, **14**, 479–481.
- Williams, M.R. & Abbott, I. (1991) Quantifying average defoliation using leaf-level measurements. *Ecology*, **72**, 1510–1511.
- Woods, E.C., Hastings, A.P., Turley, N.E., Heard, S.B. & Agrawal, A.A. (2012) Adaptive geographical clines in the growth and defense of a native plant. *Ecological Monographs*, **82**, 149–168.

Accepted 21 September 2015

First published online 20 November 2015

Associate Editor: Takayuki Ohgushi