Manipulation and Misconduct in the Handling of Image Data

The past few years have seen a small number of celebrated cases of scientific fraud that have found their way into the general media. Many more examples of inappropriate data handling come across the editorial desks of virtually every scientific journal. These have focused editors’ attention on inappropriate data handling and fraudulent image manipulation. The Plant Cell and Plant Physiology are no exceptions. Two decades ago, the practicalities of image handling meant that the boundaries were well-defined between what was acceptable and what was not; the darkroom skills needed posed a significant technical barrier to inappropriate manipulation of image data, particularly manipulation done without the intention to deceive but simply to “clean up” the image. The ethical boundaries are as clear-cut today as they were a quarter century ago, but many of the technical barriers to inappropriate manipulation have all but disappeared with the advent of digital image acquisition, storage, and handling. Adobe Photoshop was introduced in 1990 for Macintosh and in 1992 for PCs; its widespread application, and the broader acceptance of digital formats during this past decade, have simplified greatly the tasks of image preparation. They also mean that much less skill is needed to manipulate images. Indeed, a common problem arising from digital formats is that many scientists inadvertently manipulate their image data, often in ways that result in the loss of important information, to make their data look as good as possible.

The Journal of Cell Biology carried out a detailed study over the past decade and, commendably, has shared this information publicly. The study found that 10% of articles accepted for publication included inappropriate manipulations of image data that contravened journal policy, even if they did not alter the conclusions drawn from the data (International Society of Managing and Technical Editors, 2013). A surprisingly large number of the authors appeared unaware that they had handled image data inappropriately and, in many cases, were not conscious of the ethical issues and consequences of their actions. As editors, how do we maintain ethical standards in publishing? And, as scientists, how do we educate our students and support our peers to understand what is (and what is not) acceptable practice when handling image data?

It is essential to recognize that digital images are data, in fact arrays of numerical data, and must be treated as such. As scientists, we assume that images will not have been altered in any way that affects the visual impression; the quantitative and qualitative relationships within images (data arrays) must be maintained. If these relationships are altered, then such alterations must be fully documented and explained. There are two defining principles behind these expectations: (1) We expect honesty and transparency in scientific reporting, and (2) We expect the scientist, as author, to understand the consequences of processing image data to ensure that any transformations are quantitatively rigorous and comply with ethical standards. There are a few simple rules to follow in meeting these expectations (Rossner and Yamada, 2004; North, 2006; Cromey, 2010):

1. Raw image data must be saved and archived intact and without alteration as part of good laboratory practice. Processing of digital images should be done on a copy of the image data file, not on the original. Retaining raw image data is important because they serve as the standard against which the final image can be compared, and they ensure a route for recovery should a mistake be made during processing. We recommend that image data are saved in TIF format. JPEG compression affects the resolution of the image, and information is lost in the process of conversion.

2. Simple adjustments, applied uniformly, to the entire image are generally acceptable. Changes to brightness, contrast, and color balance fall into this category because they affect the image in a linear fashion. However, it is not acceptable to adjust brightness or contrast levels to such an extent that image data are truncated or lost (giving a white or black background; see Fig. 1B). Such changes may give a clearer picture of bands which are “of interest” in a gel, but they will mask background, including information that is important for quantification and validation. We will not accept image data that are processed in this way.

3. Cropping and resizing an image is usually acceptable, but both may be construed as inappropriate manipulation, on occasion. If cropping, ask whether your motivation is to improve the composition of the image or to hide something that complicates interpretation. The former reason is acceptable; the latter is not.

4. Digital filtering of an image is not encouraged because it can easily mask important information. Most filters use mathematical functions that are nonlinear. There are circumstances in which digital filtering is a necessary part of the experimental methodology. If so, filter processing must be clearly justified and documented in the figure legend or under “Materials and Methods.” Such documentation should include reference to the software version, specification of the filters, and any special settings that were used.

5. Combining images is acceptable only if it is clear to the reader that the images are from separate sources. It is acceptable to combine the images of two similar gels or two parts of the same gel in one figure, but only if a visible gap is left between the images or the images
are separated and each surrounded by a box. It is not acceptable to splice two gel images together so that they appear to be adjacent tracks from a single gel.

6. Selective alteration or processing of one region of an image is not acceptable. Such manipulations include “cloning” or copying objects or sections within or between images, “smudging,” blurring, blending, and other manipulations that are applied locally within an image. Common examples (see Fig. 1A) involve sections of an image that have been cloned or blended, to clean up a dirty preparation or to mask an unwanted blemish. Such manipulations constitute inappropriate handling at best and are unethical. If the data require such processing, repeat the experiment.

7. When comparing digital images, it is important that they be acquired under identical conditions, and any postacquisition image processing must be applied identically. If the background or color balance must be adjusted between images within a group, this must be acknowledged in the figure legend or under “Materials and Methods” (see Fig. 1B). Quantitative analysis of images should always be carried out on uniformly processed image data, and the data should be calibrated to a known standard. Most instruments, including fluorescent microscopes, are prone to fluctuations and drift over time, so it is advisable to include appropriate internal standards as checks against such changes.

8. Image data should be documented both with representative images as well as with quantitative statistical analysis of sufficient numbers of experiments. It should be self-evident that experiments that include image data should be repeated and the data analyzed for significance. We expect conclusions drawn from image data to be justified based on their quantitative assessment, not on anecdotal observations.

As editors, we have a responsibility to the readers and authors of Plant Physiology and The Plant Cell to ensure that what we publish is sound scientifically and meets the highest ethical standards. We can help authors become aware of data mishandling and the ethical consequences of inappropriate manipulations and address the probable 10% of articles falling into the category of data handling that is simply misguided or ethically ignorant. Most inappropriate data handling is relatively easy to spot and is often flagged by reviewers. From an editorial and educational standpoint, it is always best to identify and deal with such instances before an article is accepted. To this end, Plant Physiology and The Plant Cell will now have available the facility to analyze cases of suspect mishandling using the forensic tools used by The Rockefeller Press journals, including The Journal of Cell Biology. We are confident that these tools will give our editors the resources they need to handle problems of inappropriate data handling as and when questions arise. We hope, too, that our approach to these issues will help strengthen the scientific community and the reliability of the data we publish.

Mike Blatt
Editor-in-Chief
Plant Physiology

Cathie Martin
Editor-in-Chief
The Plant Cell

LITERATURE CITED


