Mauro Barni and
Fernando Perez-Gonzalez

Pushing Science into Signal Processing

Experimental analysis plays a major role in modern science. Since the time of Galileo (16th century), it has been widely recognized that any modern scientific theory or work must be backed by a strong experimental analysis carried out scrupulously and with the maximum rigor.

Replicability is one of the main requirements, possibly the most important requirement, of a well-conducted experiment. For an experiment to be credible, the experimental setup must be described very accurately so that any researcher can reproduce it and obtain the same results. If other scientists cannot reproduce your results, chances are it may have been due to a flaw in the experiment rather than a real effect. This is why great attention is paid to the experimental part in traditional scientific disciplines such as physics, chemistry, and biology. In fact, in some cases, scientists tend to specialize in theoretical or experimental scientific research.

Another important concept common in theoretical science is the use of refutation criteria; a good theoretical work should inform the reader about those events that, if observed, clearly disprove a hypothesis. This concept was clearly stated by K. Popper in his assertion that you cannot prove that a theory is correct: you can only refute it. But if you provide a reasonable set of refutation criteria and your theory survives, then you gain extra confidence in its explanatory value.

As readers of *IEEE Signal Processing Magazine,* we all love this field and think of it as a truly scientific discipline. Yet, there is much we can learn from Galilean science to improve the quality of our research, conferences, and journals. This is not the right forum to discuss whether signal processing can be treated by the same standards as true science (whatever that may be) or be left in the limbo of technology, halfway between earth and heaven. Nor do we want to be engaged in this kind of discussion. On the contrary, our goal is to point out how many times signal processing works (at least in the form in which they are published and reported to the international signal processing community) do not match one of the the basic rules of scientific thought: experimental validation.

Think about image processing, a field we are more familiar with. It is not rare, nay it is rather common, to read very nice theoretical papers validated by a single experiment carried out on a single image. It is also very common to read papers where the superiority of the proposed algorithm is claimed without any meaningful comparison with competing schemes. It is also common for an author to justify his case simply by picking up from the sea of similar algorithms an algorithm that performs worse than the proposed one, at least for the single-image, single-test experiment shown in the paper. Even worse, in many cases the experimental conditions are only vaguely described so that replicability of results is nothing but a dream. Readers must simply trust the authors, except for the noticeable case in which the results of an algorithm of yours are shown and they look much worse than you expected.

We believe that today it is almost impossible to perform research that covers both the theoretical and experimental aspects, this also being due to the increasing complexity of engineering projects as well as the *publish or perish* pressure. But from theoretical science we should at least learn how to use refutation criteria. Even if we do not have theories (at least in the common sense in which this term is used), in many cases we do have models. Many times, significant advances come from the fact that we allegedly have a better model or we have a better use for a model (as it happens with algorithms). In those cases, the problem boils down to whether the model fits the physical reality. Minor algorithmic improvements are, at this epistemological level, what Thomas Kuhn called "puzzle solving." Yet, how many theoretical papers in signal processing supply proper refutation criteria? Even though theoretical researchers can be forgiven for not having carried out the experiments, we should ask them to pinpoint which experimental outcomes would *falsify* their theory.

Who is to blame for this kind of mess? Surely signal processing researchers are not so narrow minded to not recognize the importance of well-conducted experiments. We do not think bad faith has much to do with experiment shallowness; the honor of publishing a paper does not deserve our lies. Lack of time is one of the reasons. However, it is our opinion that the main problem is the lack of research groups expressly devoted to experimental signal processing or research groups spending time and resources implementing someone else's algorithms and repeating their experiments or carrying out new experiments trying to understand the validity of different algorithms and compare them. Is this so difficult? Certainly it is not rewarding. Imagine a paper showing that the experimental results presented somewhere else are wrong. Where could such a paper be published? More likely than not, the paper would be rejected for lack of novelty. Yet, these kinds of papers would be very beneficial: they would shed some light on the effectiveness of

(continued on page 119)

1053-5888/05/$20.00©2005IEEE

 continued from page 120

popular, yet unexplored, algorithms and would encourage researchers to carefully design (and describe) their experiments to avoid the poor performance of their work being publicly revealed to the international community. Unfortunately, our journals give extremely little value to negative results, and this has done and continues to do a lot of harm.

First, think about the time that could be saved if we were told in advance that this or that idea is not going to work. Second, since you know that your experimental paper simply showing the incompetence of someone else's *algorithm A* is not going to be published, you rack your brain until you find *improvement B*, or worse yet, a very specific instance where B represents a true improvement over A. And of course, this usually entails a biased and insignificant contribution (if any) to the real progress of signal processing. Is it because no one enjoys bad news? In fact, considering the previous discussion on replicability, nothing would be more valuable than a good negative result that serves to falsify a theory. "Give me a fruitful error any time, full of seeds, bursting with its own corrections. You can keep your sterile truth for yourself," said Vilfredo Pareto.

What's the solution then? No magic wand will work here. The problem cannot be solved by acting on the review mechanism. On the contrary, let us start making experimental signal processing rewarding by creating opportunities for all-experimental, no-novelty works to be presented or published, such as by promoting special sessions at our conferences and special issues/sections in our prestigious journals. In biology, there is a (serious) publication named *Journal of Negative Results*, which publishes works that "i) test novel or established hypotheses/theories that yield negative or dissenting results, or ii) replicate work published previously (in either cognate or different systems)." Wouldn't it make sense to have a "negative results" section in our journals? We should also start teaching our students how important this kind of work is and train them to do it properly. Surely, in the end, some of the Galilean method will leak into our beloved signal processing discipline. [SP]