Ockham's Razor

Ockham's Razor and Chemistry by *Roald Hoffmann, Vladimir I. Minkin, Barry K. Carpenter*. In HYLE--An International Journal for the Philosophy of Chemistry, Vol. 3 (1997).
We begin by presenting William of Ockham's various formulations of his principle of parsimony, Ockham's Razor. We then define a reaction mechanism and tell a personal story of how Ockham's Razor entered the study of one such mechanism. A small history of methodologies related to Ockham's Razor, least action and least motion, follows. This is all done in the context of the chemical (and scientific) community's almost unthinking acceptance of the principle as heuristically valuable. Which is *not* matched, to put it mildly, by current philosophical attitudes toward Ockham's Razor. What ensues is a dialogue, pro and con. We first present a context for questioning, within chemistry, the fundamental assumption that underlies Ockham's Razor, namely that the world is simple. Then we argue that in more than one pragmatic way the Razor proves useful, without at all assuming a simple world. Ockham's Razor is an instruction in an operating manual, not a world view. Continuing the argument, we look at the multiplicity and continuity of concerted reaction mechanisms, and at principal component and Bayesian analysis (two ways in which Ockham's Razor is embedded into modern statistics). The dangers to the chemical imagination from a rigid adherence to an Ockham's Razor perspective, and the benefits of the use of this venerable and practical principle are given, we hope, their due.

Key Concepts in Model Selection:Performance and Generalizability, M. R. Forster (July 8, 1998), invited for a forthcoming special issue on model selection in the *Journal of Mathematical Psychology*.
        What is model selection?
        What are the goals of model selection?
        What are the methods of model selection, and how do they work?
        Which methods perform better than others, and in what circumstances?
These questions rest on a number of key concepts in a relatively underdeveloped field. The aim of this essay is to explain some background concepts, highlight some of the results in this special issue, and to add my own.
  The standard methods of model selection include classical hypothesis testing, maximum likelihood, Bayes method, minimum description length, cross-validation and Akaike's information criterion. They all provide an implementation of Occam's razor, in which parsimony or simplicity is balanced against goodness-of-fit. These methods primarily take account of the sampling errors in parameter estimation, although their relative success at this task depends on the circumstances. However, the aim of model selection should also include the ability of a model to generalize to predictions in a different domain. Errors of extrapolation, or generalization, are different from errors of parameter estimation. So, it seems that simplicity and parsimony may be an additional factor in managing these errors, in which case the standard methods of model selection are incomplete implementations of Occam's razor.

The New Science of Simplicity by M. R. Forster (1999): forthcoming in *Simplicity, Inference and Econometric Modelling*, Cambridge University Press, edited by Hugo Keuzenkamp, Michael McAleer, and Arnold Zellner.
There was time when statistics was a mere footnote to the methodology of science; concerned only with the mundane task of estimating the size of observational errors and designing experiments. That was because statistical methods assumed a fixed background "model", and only methodology was concerned with the selection of the model. Simplicity was an issue in methodology, but not in statistics. All that has changed. Statistics has expanded to cover model selection, and simplicity has appeared in statistics with a form and precision that it never attained in the methodology of science. This is the new science of simplicity.
  This paper lays a foundation for all forms of model selection from hypothesis testing and cross validation to the newer AIC and BIC methods that trade off simplicity and fit. These methods are evaluated with respect to a common goal of maximizing predictive accuracy. Within this framework, there is no relevant sense in which AIC is inconsistent, despite an almost universally cited claim to the contrary. Asymptotic properties are not pivotal in the comparison of selection methods. The real differences show up in *intermediate* sized data sets. Computer computations suggest that there are no global optimums—the dilemma is between performing poorly in one set of circumstance or performing poorly in another.