

features of experimental design: controls, replication, randomization, and interspersing.

One can always *assume* that certain sources of confusion are not operative and simplify experimental design and procedures accordingly. This saves much work. However, the essence of a controlled experiment is that the validity of its conclusions is *not* contingent on the concordance of such assumptions with reality.

Against the last source of confusion listed (Table 1), experimental design can offer no defense. The meaning of demonic and nondemonic intrusion will be clarified shortly.

Controls.—"Control" is another of those unfortunate terms having several meanings even within the context of experimental design. In Table 1, I use control in the most conventional sense, i.e., any treatment against which one or more other treatments is to be compared. It may be an "untreated" treatment (no imposition of an experimental variable), a "procedural" treatment (as when mice injected with saline solution are used as controls for mice injected with saline solution plus a drug), or simply a different treatment.

At least in experimentation with biological systems, controls are required primarily because biological systems exhibit temporal change. If we could be absolutely certain that a given system would be constant in its properties, over time, in the absence of an experimentally imposed treatment, then a separate control treatment would be unnecessary. Measurements on an experimental unit prior to treatment could serve as controls for measurements on the experimental unit following treatment.

In many kinds of experiments, control treatments have a second function: to allow separation of the effects of different aspects of the experimental procedure. Thus, in the mouse example above, the "saline solution only" treatment would seem to be an obligatory control. Additional controls, such as "needle insertion only" and "no treatment" may be useful in some circumstances.

A broader and perhaps more useful (though less conventional) definition of "control" would include *all* the obligatory design features listed beside "Sources of confusion" numbers 1-6 (Table 1). "Controls" (*sensu stricto*) *control* for temporal change and procedure effects. Randomization *controls* for (i.e., reduces or eliminates) potential experimenter bias in the assignment of experimental units to treatments and in the carrying out of other procedures. Replication *controls* for the stochastic factor, i.e., among-replicates variability inherent in the experimental material or introduced by the experimenter or arising from nondemonic intrusion. Interspersion *controls* for regular spatial variation in properties of the experimental units, whether this represents an initial condition or a consequence of nondemonic intrusion.

In this context it seems perfectly accurate to state that, for example, an experiment lacking replication is

TABLE 1. Potential sources of confusion in an experiment and means for minimizing their effect.

Source of confusion	Features of an experimental design that reduce or eliminate confusion
1. Temporal change	Control treatments
2. Procedure effects	Control treatments
3. Experimenter bias	Randomized assignment of experimental units to treatments Randomization in conduct of other procedures "Blind" procedures*
4. Experimenter-generated variability (random error)	Replication of treatments
5. Initial or inherent variability among experimental units	Replication of treatments Interspersion of treatments Concomitant observations
6. Nondemonic intrusion†	Replication of treatments Interspersion of treatments
7. Demonic intrusion	Eternal vigilance, exorcism, human sacrifices, etc.

* Usually employed only where measurement involves a large subjective element.

† Nondemonic intrusion is defined as the impingement of chance events on an experiment in progress.

also an uncontrolled experiment; it is not controlled for the stochastic factor. The custom of referring to replication and control as separate aspects of experimental design is so well established, however, that "control" will be used hereafter only in this narrower, conventional sense.

A third meaning of control in experimental contexts is regulation of the conditions under which the experiment is conducted. It may refer to the homogeneity of experimental units, to the precision of particular treatment procedures, or, most often, to the regulation of the physical environment in which the experiment is conducted. Thus some investigators would speak of an experiment conducted with inbred white mice in the laboratory at $25 \pm 1^\circ\text{C}$ as being "better controlled" or "more highly controlled" than an experiment conducted with wild mice in a field where temperature fluctuated between 15° and 30° . This is unfortunate usage, for the adequacy of the true controls (i.e., *control treatments*) in an experiment is independent of the degree to which the physical conditions are restricted or regulated. Nor is the validity of the experiment affected by such regulation. Nor are the results of statistical analysis modified by it; if there are no design or statistical errors, the confidence with which we can reject the null hypothesis is indicated by the value of *P* alone. These facts are little understood by many laboratory scientists.

This third meaning of control undoubtedly derives in part from misinterpretation of the ancient but ambiguous dictum, "Hold constant all variables except the one of interest." This refers not to temporal con-