and the reduction in number of cases are avoidable through the modern statistics which supplanted the matching-error in Design 10. The matching variables could all be used as covariates in a multiple-covariate analysis of covariance. It is our considered estimate that this analysis would remove the apparently significant effects in the specific studies which Chapin presents. (But see Lord, 1960, for his criticism of the analysis of covariance for such problems.) There is, however, a second and essentially uncorrectable source of undermatching in Chapin's setting. Greenwood (1945) refers to it as the fact of self-selection of exposure or nonexposure. Exposure is a lawful product of numerous antecedents. In the case of dropping out of high school before completion, we know that there are innumerable determinants beyond the six upon which matching was done. We can with great assurance surmise that most of these will have a similar effect upon later success, independently of their effect through $X$. This insures that there will be undermatching over and above the matching-regression effect. Even with the pre-$X$-predictor and $O$ covariance analysis, a significant treatment effect is interpretable only when *all* of the jointly contributing matching variables have been included.

## CONCLUDING REMARKS

Since a handbook chapter is already a condensed treatment, further condensation is apt to prove misleading. In this regard, a final word of caution is needed about the tendency to use the speciously convenient Tables 1, 2, and 3 for this purpose. These tables have added a degree of order to the chapter as a recurrent outline and have made it possible for the text to be less repetitious than it would otherwise have been. But the placing of specific pluses and minuses and question marks has been continually equivocal and usually an inadequate summary of the corresponding discussion. For any specific execution of a design, the check-off row would probably be different from the corresponding row in the table. Note, for example, that the tie-breaking case of Design 6 discussed incidentally in connection with quasi-experimental Design 16 has, according to that discussion, two question marks and one minus not appearing in the Design 6 row of Table 1. The tables are better used as an outline for a conscientious scrutiny of the specific details of an experiment while planning it. Similarly, this chapter is not intended to substitute a dogma of *the* 13 acceptable designs for an earlier dogma of *the* one or *the* two acceptable. Rather, it should encourage an open-minded and exploratory orientation to novel data-collection arrangements and a new scrutiny of some of the weaknesses that accompany routine utilizations of the traditional ones.

In conclusion, in this chapter we have discussed alternatives in the arrangement or design of experiments, with particular regard to the problems of control of extraneous variables and threats to validity. A distinction has been made between internal validity and external validity, or generalizability. Eight classes of threats to internal validity and four factors jeopardizing external validity have been employed to evaluate 16 experimental designs and some variations on them. Three of these designs have been classified as pre-experimental and have been employed primarily to illustrate the validity factors needing control. Three designs have been classified as "true" experimental designs. Ten designs have been classified as quasi-experiments lacking optimal control but worth undertaking where better designs are impossible. In interpreting the results of such experiments, the check list of validity factors becomes particularly important. Throughout, attention has been called to the possibility of creatively utilizing the idiosyncratic features of any specific research situation in designing unique tests of causal hypotheses.

## REFERENCES

Allport, F. H. The influence of the group upon association and thought. *J. exp. Psychol.*, 1920, 3, 159–182.