Eta-Squared and Partial Eta-Squared in Fixed Factor Anova Designs

Jacob Cohen

Educational and Psychological Measurement 1973 33: 107
DOI: 10.1177/001316447303300111

The online version of this article can be found at:
http://epm.sagepub.com/content/33/1/107

Published by:
SAGE
http://www.sagepublications.com

Additional services and information for Educational and Psychological Measurement can be found at:

Email Alerts: http://epm.sagepub.com/cgi/alerts
Subscriptions: http://epm.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav
Citations: http://epm.sagepub.com/content/33/1/107.refs.html
ETA-SQUARED AND PARTIAL ETA-SQUARED IN FIXED FACTOR ANOVA DESIGNS

JACOB COHEN
New York University

In a recent issue of this journal, Kennedy (1970) distinguished between what he identified as two formulas for $\eta^2$ (eta-squared) to be used in fixed model ANOVA. The first he called "the classical formula proposed by Kerlinger (1964, pp. 200–206),"

$$\eta^2 = \frac{SS_A}{SS_T},$$

(1)

where $SS_A$ is the between sum of squares for factor $A$, and $SS_T$ is the total sum of squares. $\eta^2$ is simply the proportion of the total $SS$ (or variance) associated with $A$. The second he identified as "an alternate version of the $\eta^2$ computational formula which is not confined to the single independent variable case" proposed by Cohen (1965) and further described by Friedman (1968),

$$\eta^2 = \frac{df_A F_A}{df_A F_A + df_B},$$

(2)

where $df_A$ and $df_B$ are the degrees of freedom for between $A$ groups and error, respectively, the $F_A$ value being the one for the $A$ effect.\(^1\)

Kennedy goes on to point out, quite correctly, that these two formulas, which are algebraically identical when there is only one research factor, cease to be so when $A$ is one of two or more factors in a complex design. This is so because the Cohen formula

---

\(^1\) Kennedy mistakenly adds the qualification to this formula "when $F < 1."$ Since the relationship is purely algebraic, the formula is in fact valid for all possible values of $F$, although when $F < 1$, epsilon-squared is negative and customarily reported as zero.

107
(2) was derived from

$$\eta^2 = \frac{SS_A}{SS_A + SS_E},$$

(3)

and in a one-way ANOVA,

$$SS_T = SS_A + SS_E,$$

(4)

making the equivalent formulas (3) and (2) also equivalent to formula (1). But in an experiment with, for example, three identified sources of variance, obviously

$$SS_T = SS_A + SS_B + SS_C + SS_E,$$

(5)

and formula (1) gives for factor A's proportion of variance

$$\eta^2 = \frac{SS_A}{SS_A + SS_B + SS_C + SS_E},$$

(6)

which is obviously a different (and necessarily smaller) quantity than given by formula (3) and hence by the equivalent Cohen formula (2): formulas (1) and (6) give factor A's proportion of total SS or variance, while formulas (2) and more obviously (3) give factor A's proportion of total less that due to factors B and C.

It is thus fundamentally an error to consider formulas (1) and (2) as "alternative." Put succinctly, formula (1) is quite properly the formula for \( \eta^2 \), while formula (2) is a formula for partial \( \eta^2 \), with "other nonerror sources of variance being partialled out" (Cohen, 1965, p. 105). The term "partial" is not being used analogically but quite literally and in exactly the same sense as it would be in connection with product-moment correlation coefficients. The reason for the equivalence of formulas (1) and (2) for the one-factor design is that in the absence of other factors, formula (2) partials nothing out and reduces to formula (1), just as partial \( r \) reduces to \( r \) when there are no variables being partialled.

Thus, formula (2) must be rewritten for the record so that it is unambiguously identified as the partial \( \eta^2 \) of a research factor A with a dependent variable Y from which all other nonerror sources of variance (main effects, interactions, trend components, etc.) in the experiment (B, C, \( \cdots \) J) have been removed:

$$\eta_{YA \cdot BC \cdots J}^2 = \frac{df_AF_A}{df_AF_A + df_E} = \frac{SS_A}{SS_A + SS_E}.$$

(7)
The burden of Kennedy's article is his preference for $\eta^2$ over partial $\eta^2$. This is maintained on two grounds, one of which (clearly of lesser importance to him) is that in a complex fixed effects design, the $\eta^2$'s of all the effects can be assessed in terms of their contribution to the total variance, and are additive to the total proportion of "explained" variance. Partial $\eta^2$ values are, like all squared partial coefficients, not additive since they do not share a common base. This may make the unpartialled $\eta^2$ advantageous for some descriptive purposes, but, as will be demonstrated shortly, if the nature of the application renders $\eta^2$ a fundamentally inappropriate measure of effect size for the factors under study, the property of additivity is hardly a compensating consideration.

Let us then turn to Kennedy's major point, which is that partial $\eta^2$ for a given constant research factor $A$, will vary from experiment to experiment depending on the number and combined potency of other factors $B$, $C$, ..., $J$, while $\eta^2$ will not. The apparent validity of this argument rests on comparing the denominators of the formula (3) version of partial $\eta^2$ with the formula (6) version of $\eta^2$ and assuming that in an experiment in which, for example, only $A$ is studied, variables $B$ and $C$ nevertheless operate to increase the total variance and hence the error variance. The other side of this assumption is that in the experiment which studies $B$ and $C$ as well as $A$, the error variance will be reduced by the removal of $B$ and $C$ variance, thus affecting partial $\eta^2$, but not $\eta^2$, in the comparison of such experiments. That this view is invalid for many experiments is readily demonstrated.

**Illustrative Example**

Consider an educational experiment in which teaching methods effects ($A$) on an achievement criterion are to be studied in a one-way design with randomized assignment of pupils. Good craftsmanship in experimentation would dictate that the teachers used would be of comparable level of experience and the pupils would be relatively homogeneous in demographic background (socioeconomic status, ethnicity). Dr. Doe performs the research and finds that $SS_A = 100$, $SS_E = 900$, and therefore $SS_T = 1000$.

---

2 Kennedy mistakenly states (p. 889) that "the sum of all eta-squared values equals unity," apparently a slip of the pen.
Note that $SS_B$ contains no effects of variation in teacher experience or pupil demography. Thus, $\eta^2$ of formula (1) gives $100/1000 = .1000$, as does the partial $\eta^2$ of formula (7): $100/(100 + 900) = .1000$, there being no other effect to partial, as we have already seen.

Now consider Dr. Roe, who performs an experiment which uses the same teaching methods factor ($A$) as did Dr. Doe, but also studies amount of teaching experience ($B$), and the interaction of these two factors as a third potential source of variability ($C$). Given an experiment of the same size as the other, he would have as expected $SS : SS_A = 100$ and $SS_B = 900$ as before, and assume that he finds for the new sources: $SS_B = 200$ and $SS_C = 40$, the new $SS_T$ therefore being the sum of all these values, 1240. When Dr. Roe computes $\eta^2$ from formula (1) it is $100/1240 = .0806$, obviously changed from Dr. Doe's result, while the partial $\eta^2_{BA, BC}$ formula (7), now however working to remove $B$ and $C$ variance from the total, gives, as before, $100/(100 + 900) = .1000$. A third experimenter adding socioeconomic status ($D$) and ethnicity ($E$), and all the resulting interactions (now a four-factor design with a total of $2^4 - 1 = 15$ nonerror sources of variance) none of which operated in the $SS_B$ of the first experiment, has the same partial $\eta^2$ expectation of .1000 for $A$, but an $\eta^2$ further diminished to a degree that depends upon the total variance due to these many additional effects, all now swelling the base.

Discussion

Kennedy does not make sufficiently clear the basis of his strictures against the use of partial $\eta^2$. The above example makes clear that there are circumstances where sources of variance in some given Experiment II should be partialled out of total (and, therefore, out of error) in order to make the proportion of variance for some source of interest in Experiment II comparable to that for the same source in a reference Experiment I. In general, one wishes to partial out variables which are not properly considered "error" in one's conception of the background against which a factor $A$ is to be appraised. More specifically, at least the following two kinds of variables are obvious candidates for partialling:

a. Manipulated variables. A variable which exists because of
systematic manipulation by an experimenter must be removed if the effect size of A is to be compared with that obtained in some reference experiment, since by definition the manipulated variable accounts for none of the SS in the reference experiment.

b. Variables controlled by being held constant. The above example illustrates variables of this kind. Dr. Doc, who was studying only methods, held constant teacher experience and pupil socioeconomic status and ethnicity. These variables, therefore, did not inflate SS in his experiment, thus making it important to remove them from the SS in the other experiments where they were part of SS.

Variables which should not be partialled in an Experiment II are those which had contributed to SS in the reference Experiment I. The circumstance Kennedy has in mind (p. 888) is one in which later experiments remove variance due to sources which had operated in earlier experiments by introducing them explicitly as control factors. Only in such cases is it appropriate to use $\eta^2$.

Cases will arise in which the A factor occurs in an experiment in which one or more of the other factors explicitly under study should be partialled and one or more of them should not. In such cases neither $\eta^2$ nor the partial $\eta^2$ of formula (7) is appropriate, since the latter partials all other sources out. One then needs a custom-tailored partial $\eta^2$ in which the denominator includes the SS for those sources which operated in the reference experiment, and excludes those which did not. Thus, in the illustrative example, the first experimenter held teacher experience (B) constant so that neither it nor its interactions with other variables could produce any variability. But assume that he used pupils who were heterogeneous in regard to socioeconomic status (D) and ethnicity (E). Then his one-factor study of A would have had its SS inflated by D, E, D x E, and also by A x D, A x E, and A x D x E. The third experimenter, explicitly studying all these factors, would then want to include all these SS in the denominator in comparing his A effect with that of the first experiment, but not the SS due to B nor all the interactions involving B.

It should be noted that the above logic extends directly to fixed-factor covariance designs. When Experiment I uses a one-way covariance design, the $\eta^2$ of formula (1) applied to “adjust” SS is a partial $\eta^2$, since the nature of the adjustment is the partialling
of the covariate(s). If then, Experiment II is a complex ANOVA which uses the covariates as explicit design factors, it is the partial \( \eta^2 \) of formula (7) which is appropriately used for comparison. Incidentally, formula (7) may be applied to either the \( F \)-values or "adjusted" SS of a complex covariance design, with the resulting partial \( \eta^2 \) not only partialling all the other factors, interactions, trend components, or whatever, but also the covariates, the latter having been removed by the "adjustment."

The conclusion is banal: in data analysis, one cannot offer copybook maxims to stand in the place of a clear understanding of what one is doing. In the recent efforts to raise psychology's awareness of effect sizes and their measurement by the present author (Cohen, 1965, 1969) and others (Hays, 1963; Kerlinger, 1964; Friedman, 1968), proportion of variance measures have a centrally important role (Cohen, 1968, 1969). But proportions of variance, or for that matter, of apples, can only be interpreted if one understands the base of the proportion. The proportions of variance defined by \( \eta^2 \) and partial \( \eta^2 \) refer to different bases, and the choice between them and their interpretation demands an understanding of the sources which have contributed to the base, and particularly to \( SS_{\eta} \).

REFERENCES


