A TEST OF DURKHEIM'S THEORY OF SUICIDE— WITHOUT COMMITTING THE "ECOLOGICAL FALLACY"*

Frans van Poppel
Netherlands Interdisciplinary Demographic Institute

Lincoln H. Day
Washington, D.C.

The data adduced by Durkheim in support of the association between religion and suicide have seldom been subjected to scrutiny; when they have been so examined, the scrutiny has been based, of necessity, on data subject to the "ecological fallacy." Data for the Netherlands, roughly contemporaneous with Durkheim's, that have recently come to light allow us to test the statistical support for Durkheim's theory about religion and suicide without risk of committing this "fallacy." We find the Catholic-Protestant differential in suicide rates to be explicable entirely in terms of the practice of categorizing as "sudden deaths" or "deaths from ill-defined or unspecified cause" a large proportion of deaths among Catholics which would have been categorized as suicides had they occurred among Protestants. This finding raises doubts not only about Durkheim's theory but also about other causal theories concerning suicide that rely on a sociological rather than a psychological (or even idiosyncratic) explanation.

Almost a century has passed since the publication of Durkheim's ([1897] 1951) *Suicide*. Although his is not the first sociological analysis of suicide—that honor probably belongs to the work of Masaryk ([1881] 1970)—the analytical rigor and theoretical underpinning of *Suicide* have made it the most influential of works (e.g., see Merton 1967). Even today, it is the customary starting point for both the sociological and the epidemiological analysis of suicide.

Most of the considerable scholarly attention accorded *Suicide* has focused on Durkheim's discussion of the association between religion and suicide, particularly the higher suicide rates in Protestant than in Catholic populations. Durkheim's analysis of this association rests on two assumptions: first, that his statistical data contain essentially no religion-based biases; and second, that Protestant and Catholic societies differ from one another in ways of causal significance to suicidal behavior. Durkheim paid only scant attention to the first assumption: His one expression of doubt on this score was relegated to a mere footnote (Durkheim [1897] 1951:160). About the second assumption, however, which is the theoretical meat of his discussion, he was much more expansive—and also more creative. Rather than accounting for the observed differences between Protestant and Catholic suicide rates in terms of the theological and dogmatic differences that separate the two denominations, Durkheim created an explanation in terms of the respective *social* differences he presumed to be indicated by Protestantism and Catholicism.

Suicide has not lacked critics. These have confined themselves mostly to the theory and method of the study (Selvin 1957–1958, 1965; Pope 1976; Pope and Danigelis 1981; Stark, Doyle, and Rushing 1983), and few criticisms have been directed at the statistical materials Durkheim adduced in support of his argument. Like Durkheim himself, the critics of *Suicide* have been notably accepting of the data Durkheim used (e.g., Pescosolido and Georgianna 1989).

Yet no comparison can be reliable if the measures employed do not mean essentially the same thing for each population under

* Address correspondence to Frans van Poppel, NIDI, P.O. Box 11650, 2502 AR The Hague, Netherlands (poppel@nidi.nl).

consideration. Nor can any theory—however elegant or appealing—be of more than questionable validity if formulated in explanation of a nonexistent difference. If rates for a particular phenomenon in one population are based on definitions or recording practices that differ from those in another population, comparisons will be biased and any claim of a difference between the two will be flawed.

More than three decades ago, Kruijt (1960) suggested that the differences in suicide rates recorded by religious composition in the Netherlands might be explicable largely in terms of nothing more arcane than differences in recording practices:

> [T]he lowest registration of suicides . . . in the south of the Netherlands, and perhaps also in some Orthodox Protestant areas, may correlate with prevailing negative attitudes toward filling in "suicide" as a cause of death on the death certificate. There is a stronger tendency among Roman Catholics to conceal the fact that suicide was the cause of death than among groups that do not judge suicide so harshly or punish suicide with religious sanctions. By pretending ignorance or even by blatantly telling the family doctor that the death was accidental, the relatives try to avoid the scandal of having the deceased buried in unconsecrated soil. This is an understandable reaction among groups with extremely high social control who seem to fear the judgment of the priest, neighbors, and acquaintances even more than they fear the wrath of God. . . .
>
> However, it is suspected that, considering the extreme aversion to suicide and its unpleasant consequences, it will likely occur more often that the principle of "when in doubt, do nothing" will be followed [by the physician filling in the cause of death on the death certificate]. If the relatives themselves are in doubt whether or not the deceased wished to die, making it either suicide or an accident, then it is obvious that the actual circumstances are no longer very relevant, as there is nothing more to gain but a lot to lose! . . .
>
> The family doctor or physician who signs the death certificate also plays a definite role. . . . In cases where the diagnosis is difficult on purely medical grounds, the Roman Catholic doctor will be sooner inclined to "turn a blind eye," particularly as he is well aware of what the diagnosis "suicide" would imply in a Roman Catholic milieu where the entire family is already suffering enough. . . . (Pp. 43–45)

A few years later, Douglas (1967, chap. 12) not only questioned the validity of Durkheim's data on suicide but also maintained that whatever the association under consideration—whether with religion or with something else—"contemporary sociologists (especially Americans) who have used the official records on suicide [have] rarely . . . explicitly considered their value as evidence" (1968:377–81), generally assuming instead that there are "no . . . consistent biases in the data" (Douglas 1968:379). Atkinson (1978) made the same point a decade later in a similarly impressive discussion.

Day (1987), however, is the only one to have acted on this skepticism by addressing Durkheim's thesis with data similar to those used by Durkheim and relating to the same historical period. Like Kruijt (1960) before him, Day (1987:452) pointed out that whatever else might be said about the association between religion and suicide, certain elements of belief—and particularly of practice—relative to suicide, specifically in Catholicism, are likely to bias official suicide rates so severely as to prohibit their reasonable use for comparisons between Catholic and non-Catholic populations. On the basis of an analysis extended to include deaths attributed not only to suicide but also to accidents, "sudden death," and "unknown causes," in Prussian *Regierungsbezirken* and cities in 1865 to 1910, Swiss cantons and cities in 1878 to 1912, and Netherlands provinces in 1869 to 1871 and 1901 to 1904, Day (1987) concluded:

> Whether Catholic societies were at this time more "cohesive" than Protestant societies we cannot say. Nor can we affirm or deny Durkheim's views about the causation of suicide rates. Given the findings reported here and the fact that those who commit suicide are such a tiny fraction of any population that is more or less normally structured demographically, one is certainly tempted to ask whether the causes of suicide itself might not lie less in the social than in the personal, psychological—even idiosyncratic—realm, after all. (Pp. 459–60)

Sociologists apparently have not found either Douglas's (1967, 1968) or Atkinson's (1978) counsels particularly persuasive. Nor have they strayed far from the Durkheimian tradition of accounting for suicide in terms

of essentially social factors. Pescosolido and Mendelsohn (1986), for example, on the basis of an impressively detailed analysis of county data for the United States, concluded not only that social factors are of prime importance in the causation of suicide but also that their importance might well "have been underestimated by the use of official rates in empirical research" (p. 95). And Stack (1983), in a comparative analysis of 25 countries, creatively refined Durkheim's thesis to arrive at the finding that lower suicide rates—for women, but not for men—are associated not with religiously based *social integration* (as Durkheim and others have argued) but with *commitment* to a few "life-preserving" religious beliefs, values, and practices.

A major limitation in analyses such as these, and in those of Day and Durkheim as well, is the "ecological fallacy": treating group data as though they were individual data (Robinson 1950:351–57; Alker 1969). At least theoretically, suicides in a given population could take place solely among those in a religious minority, among (say) Protestants, where Protestants were in the minority, and among Catholics, where Catholics were in the minority. If the propensity to commit suicide were actually the same among this population's Protestants as among its Catholics, the distribution of suicide rates in the various sectors within this population would be U-shaped—low in those sectors where the proportion Catholic was high (and the proportion in the minority non-Catholic population available to commit suicide was correspondingly low), high where proportions Catholic and Protestant were almost equal (and thus where the minority available to commit suicide, whether Protestant or Catholic, was large), and low again where the proportion non-Catholic was high (and where, in this case, the proportion in the minority Catholic population available to commit suicide was correspondingly low).

The nature of the data available to these researchers restricted them to the examination of death by the religious compositions of the *districts* in which deaths occurred (or in which the deceased had lived), not by the religion of the deceased *individuals. In fact,* some of Durkheim's contemporaries published suicide rates for specific religions. Masaryk ([1881] 1970:88–90) referred, for example, to papers by Morselli, Wagner, and Legoyt, published respectively in 1879, 1864, and 1844. In none of these studies, however, were the cause-of-death data based on medical certification, and in none did the authors employ any form of age standardization. Were these data all that were available, the case on suicide and religion still would remain open.

THE PRESENT ANALYSIS

We suggest, however, that the following analysis goes a long way toward resolving that case; it is based on death rates calculated from data on cause of death recorded simultaneously with the deceased individual's sex, age, and (particularly significant for present purposes) religious affiliation. Unfortunately, nothing in these data concerns depth of religious feeling or the extent of religious practice, however important these might be to the determination of morbidity and the timing of death (Stack 1983; Idler and Kasl 1992). Yet at least these data do not entail the risk of committing the "ecological fallacy."

THE DATA

These data pertain to the Netherlands for 1905 to 1910. They come from the annual *Statistiek van de sterfte naar den leeftijd en naar de oorzaken van den dood* (Mortality Statistics by Age and Cause of Death) published by the Netherlands Central Bureau of Statistics (NCBS) in 1906, 1908, and 1910 and are based on data initially provided by the municipal vital registration officers to the provincial public health inspectors. Before turning any record over to the NCBS, the public health inspectors at that time were responsible for determining the appropriate cause of death from among those in the International List of Causes of Death (ICD) and incorporating the code for that cause into the record. The procedures underlying these data were the culmination of four decades of improvement in the vital registration procedures in the Netherlands, particularly with respect to causes of death. In those four decades, the quality of the universal cause-of-death reporting was enhanced through introduction (in 1865) of the requirement of com-

petent medical certification overall and, in the special case of a suspected death from violence, the requirement of a coroner's inquest: The NCBS assumed responsibility for the final processing and publication of statistics on cause of death, and the causes-of-death nomenclature was made to conform with that in the international standard developed by Bertillon (1903). (For a detailed overview of the history of cause-of-death registration in the Netherlands, see van Poppel and van Dijk 1994.)

Each data record used in the present analysis contains the cause and date of death, the address at which the death occurred, the registration number of the death certificate, and personal characteristics of the deceased, such as sex, age, marital status, occupation, and religion. We have been unable to ascertain why religion should have been coded only during 1905 through 1910 and not recorded before or after this brief period. We found no answer to this question in the annual reports of either the NCBS or the Central Statistical Committee, which functioned as an advisor for the NCBS and monitored its activities. Whatever the reasons, however, these data present a unique opportunity to test the statistical support for Durkheim's theory about religion and suicide, using data essentially contemporaneous with those used by Durkheim himself.

The limitations in these data are either of no particular consequence to the present analysis (such as the fact that information on religion recorded on the individual death record comes from the municipal population register, while that on the religious composition of the population comes from the census) or precisely the type of evidence necessary to test the points under consideration (such as the extent of incompleteness in the recording of suicide).

The deaths on which these death rates were calculated were classified by cause according to the first edition of the International List of Causes of Death (Bertillon 1903) (see Table 1) .

Table 1. Causes of Death with Codes from the International List of Causes of Death, 1903

Code	Cause of Death
155	Suicide by poison
156	Suicide by asphyxia
157	Suicide by hanging, strangulation, or garroting
158	Suicide by drowning
159	Suicide by firearms
160	Suicide by cutting instruments
161	Suicide by precipitation from a height
162	Suicide by crushing
163	Other suicides
164	Fractures (cause not specified)
165	Sprains, dislocations
166	Other accidental injuries
167	Burning by fire
168	Burning by corrosive substances
169	Sunstroke
170	Freezing
171	Electric shock
172	Drowning
173	Hunger
174	Inhalation of noxious gases
175	Other accidental poisoning
176	Other external violence
178	Sudden death
179	Unspecified or ill-defined causes of death

Source: Bertillon (1903).

PREPARATION OF THE DATA

We standardized for age with the direct method, using directly standardized mortality ratios or so-called comparative mortality figures (CMFs) (Kitagawa 1964). Although the indirect method, the calculation of standardized mortality ratios (SMRs), generally yields more precise results (that is, smaller confidence intervals), it does not lend itself to such reliable comparisons between populations with large differences in age structure (Veling and Sturmans 1981:276–78) and during the period under consideration, Protestants and Catholics differed substantially in age structure (see Netherlands Central Bureau of Statistics 1907:xli–xlii).

The directly standardized mortality rates employed in our analysis are the expected numbers of deaths per 100,000 for a specific cause of death that would have occurred in

the standard population (the population that is used as reference in the comparison) if it had experienced the cause- and age-specific death rates of the index population (the population with which we compare our standard):

$$\frac{100{,}000}{\sum(P_x)}\sum\left(P_x \cdot R_{ijx}\right),$$

in which Px equals the number of persons at age x in the standard population, and R_{ijx} equals the cause- and age-specific death rates for cause j and age x in the index population i.

The comparative mortality figures (CMFs) are the ratios between the directly standardized mortality rates for given causes of death and the crude death rates for these causes:

$$\frac{\sum\left(P_x \cdot R_{ijx}\right)}{\sum\left(P_x \cdot R_{jx}\right)} = \frac{\sum\left(P_x \cdot R_{ijx}\right)}{\sum\left(D_{jx}\right)},$$

in which R_{jx} equals the cause- and age-specific death rate for cause j and age x in the standard population, and D_{jx} equals the number of deaths for cause j at age x in the standard population.

The published mortality tabulations by religion used the following age groupings: younger than 1, 1–4, 5–13, 14–19, 20–29, 30–39, 40–49, 50–64, 65–79, and 80 and older; the age groupings used in the tabulations of the population by religion were younger than 10, 10–19, 20–29, . . . , 70–79, and 80 and older. Because the two sets of age groupings did not coincide, we regrouped them into the following categories: younger than 20, 20–29, 30–39, 40–49, 50–79, and 80 and older.

We used the Catholic population as the index and the Protestant population as the standard. The population for both groups is the average of the populations at the two census dates, December 31, 1899 and December 31, 1909, multiplied by a factor of 6.

To calculate the standard error of the CMF, we used the following formula (Breslow and Day 1987):

$$SE(\log CMF) = \frac{\left[\sum\left(P_x^2 \cdot D_{ijx}/P_{ix}^2\right)\right]^{1/2}}{\sum\left(P_x \cdot D_{ijx}/P_{ix}\right)},$$

in which D_{ijx} equals the number of deaths for cause j at age x in index population i, and P_{ix} equals the number of persons at age x in index population i.

We used the standard errors to calculate the 95 percent confidence intervals:

$$\exp^{[\log(CMF)\pm 1.96\cdot SE(\log CMF)]}.$$

We calculated the confidence intervals using logarithmic transformation. This procedure yielded asymmetric confidence intervals: The lower the CMF, the greater the asymmetry. We chose a method that would yield asymmetric results so as to ensure that the lower limit of these intervals, however low the CMF, would always exceed zero.

RESULTS

The results of these calculations for the 1905 through 1910 period are displayed in Table 2, which presents CMFs for the Protestant and the Catholic populations of the Netherlands (the "standard" and the "index" populations, respectively, in the CMF calculations) for a number of cause-of-death categories, together with their respective standard errors and 95 percent confidence intervals.

Although the age-standardized death rate from all causes during this period was 16 percent higher among Catholics (both males and females) than among their Protestant counterparts, the comparable rate for suicide was significantly lower: only 47 percent as high for Catholic males and 35 percent as high for Catholic females. A more fully detailed analysis (not shown here) found lower rates for Catholics than for Protestants for all forms of suicide, except crushing among males (where the numbers were so small that the difference could reasonably be attributed to chance). With regard to accidental death, the standardized rates among Protestants were higher for the total of accidental causes, but lower for the much less numerous "real" accidental causes. Subtracting the latter from the total of accidental causes leaves us with a rather diffuse group of external causes that might well include some actual suicides. Here again, the Catholic rate was significantly lower than the Protestant.

The relationship is reversed, however, when we consider the highly ambiguous classifications of "sudden death" (mors subita) and "cause of death unknown or unspecified"—both of which could be expected to serve as alternative classifications for sui-

Table 2. Comparative Mortality Figures for External and Ill-Defined Causes of Death, by Religion and Sex: The Netherlands, 1905–1910

Cause of Death	Religion	Comparative Mortality Figure		Standard Error		95% Confidence Interval	
		Males	Females	Males	Females	Males	Females
All suicides (155–163)	Protestant	1.00	1.00	—	—	—	—
	Roman Catholic	.47	.35	.05	.12	(.42–.52)	(.27–.44)
All accidental deaths (164–176)	Protestant	1.00	1.00	—	—	—	—
	Roman Catholic	.91	.86	.02	.03	(.88–.95)	(.81–.92)
"Real" accidents[a] (167, 169, 170, 173)	Protestant	1.00	1.00	—	—	—	—
	Roman Catholic	1.13	1.15	.06	.06	(1.00–1.27)	(1.02–1.29)
Accidental deaths minus "real" accidents (164–166, 168, 172, 174–176)	Protestant	1.00	1.00	—	—	—	—
	Roman Catholic	.89	.78	.02	.04	(.85–.93)	(.72–.84)
Sudden deaths (178)	Protestant	1.00	1.00	—	—	—	—
	Roman Catholic	1.43	1.49	.05	.05	(1.31–1.57)	(1.36–1.65)
Ill-defined or unspecified causes of death (179)	Protestant	1.00	1.00	—	—	—	—
	Roman Catholic	1.93	1.90	.01	.01	(1.89–1.97)	(1.86–1.94)
All external causes, plus sudden deaths and ill-defined causes	Protestant	1.00	1.00	—	—	—	—
	Roman Catholic	1.44	1.62	.01	.01	(1.41–1.46)	(1.59–1.65)
Deaths from all causes	Protestant	1.00	1.00	—	—	—	—
	Roman Catholic	1.16	1.16	.00	.00	(1.16–1.17)	(1.15–1.17)

[a] Accidental deaths that can reasonably be presumed to contain no "hidden" suicides.

cide. For both males and females, the Catholic rates for "sudden death" were nearly half again as high as the Protestant, and those due to "unknown or unspecified" causes were nearly twice as high. For the combined total of deaths in all the categories to which suicides might be assigned—suicide, accident, sudden death, and ill-defined or unspecified cause—the relative rates were 44 percent higher among Catholic males and 62 percent higher among Catholic females.

Thus we find a strong suggestion that for Catholics the "sudden death" and "ill-defined or unspecified cause" classifications were used as alternative classifications for suicide. Although it is not beyond reason to suppose that some Protestant suicides at that time might also have found their way into these two classifications, the slight difference between Protestant and Catholic death rates from accidents (the only other alternative to which one could assign a suicide) shows that this could not have occurred very often.

CONCLUSION

The gap between Protestant and Catholic suicide rates in the Netherlands during the years 1905 through 1910 appears to be the result of nothing more mysterious than differences in how deaths to Catholics and deaths to Protestants were recorded: A large proportion of deaths to Catholics, which would have been categorized as suicides had they

occurred to Protestants, were categorized as "sudden death" or "death from ill-defined or unspecified cause." We cannot say, on the basis of this analysis, whether Durkheim's—or some other—sociological explanation of suicide is valid. We *can* say that a sociological explanation receives no support from these data: The data, although roughly contemporary with and similar to those used by Durkheim, are far superior to his because they are not subject to the risk of committing the "ecological fallacy." If suicide is to be explained in essentially sociological rather than psychological or idiosyncratic terms, the data must not be subject to such a "fallacy."

Frans van Poppel *is Senior Research Fellow at the Netherlands Interdisciplinary Demographic Institute (NIDI), The Hague, Netherlands. His research interests include differences in contemporary morbidity and mortality by marital status and sex; historical research on marriage, divorce, and widowhood in nineteenth-century Europe; patterns of Jewish infant mortality and fertility in the nineteenth century; and the historical development of the kinship network.*

Lincoln H. Day*, before his recent retirement, was Senior Fellow in Demography at the Research School of Social Sciences, Institute of Advanced Studies, The Australian National University, Canberra, ACT, Australia. He collaborated on this paper while a Hofstee Fellow at NIDI, May to September 1994.*

REFERENCES

Alker, Hayward R. 1969. "A Typology of Ecological Fallacies." Pp. 69–86 in *Quantitative Ecological Analysis in the Social Sciences*, edited by M. Dogan and S. Rokkan. Cambridge, MA: Massachusetts Institute of Technology Press.

Atkinson, Jo M. 1978. *Discovering Suicide: Studies in the Social Organization of Sudden Death.* London, England: Macmillan.

Bertillon, Jacques. 1903. *Nomenclature des maladies arrêtées par la Commission internationale chargée de reviser les nomenclatures nosdologiques (18–21 août 1900) pour être en usage à partir de 1er janvier 1901 avec notices et annexes* (International List of Causes of Death). 1st ed. Montévain, France. Imprimerie typographique de l'école d'Alembert.

Breslow, Norman E. and N. E. Day. 1987. *Statistical Methods in Cancer Research.* Vol. 2, *The Design and Analysis of Cohort Studies.* Lyon, France: International Agency for Research on Cancer (IARC).

Day, Lincoln H. 1987. "Durkheim on Religion and Suicide—A Demographic Critique." *Sociology* 21:449–61.

Douglas, Jack D. 1967. *The Social Meanings of Suicide.* Princeton, NJ: Princeton University Press.

Douglas, Jack D. 1968. "Suicide: Social Aspects." Pp. 375–85 in *International Encyclopedia of the Social Sciences*, vol. 15. New York: Macmillan and Free Press.

Durkheim, Émile. [1897] 1951. *Suicide.* New York: Free Press.

Idler, Ellen L. and Stanislav V. Kasl. 1992. "Religion, Disability, Depression, and the Timing of Death." *American Journal of Sociology* 97:1052–1109.

Kitagawa, Evelyn M. 1964. "Standardized Comparisons in Population Research." *Demography* 1:296–315.

Kruijt, Cornelis S. 1960. *Zelfmoord: Statistisch-Sociologische Verkenningen* (Suicide: Statistical and Sociological Analysis). Assen, Netherlands: Van Gorcum.

Masaryk, Thomas. [1881] 1970. *Suicide and the Meaning of Civilization.* Chicago, IL: University of Chicago Press.

Merton, Robert K. 1967. *On Theoretical Sociology: Five Essays, Old and New.* New York: Free Press.

Netherlands Central Bureau of Statistics. 1907. *Statistiek van de sterfte naar den leeftijd en naar de oorzaken van den dood over het jaar 1906* (Mortality by Age and Cause of Death in 1906). The Hague, Netherlands: Gebrs. Belinfante.

Pescosolido, Bernice A. and Sharon Georgianna. 1989. "Durkheim, Suicide, and Religion: Toward a Network Theory of Suicide." *American Sociological Review* 54:33–48.

Pescosolido, Bernice A. and Robert Mendelsohn. 1986. "Social Causation or Social Construction of Suicide? An Investigation into the Social Organization of Official Rates." *American Sociological Review* 51:80–100.

Pope, Whitney. 1976. *Durkheim's Suicide: A Classic Analyzed.* Chicago, IL: University of Chicago Press.

Pope, Whitney and Nicholas Danigelis. 1981. "Society's 'One Law.'" *Social Forces* 60:495–516.

Robinson, William S. 1950. "Ecological Correlations and the Behavior of Individuals." *American Sociological Review* 15:351–57.

Selvin, Hanan. 1957–1958. "Durkheim's *Suicide* and Problems of Empirical Research." *American Journal of Sociology* 63:607–19.

Selvin, Hanan. 1965. "Durkheim's Suicide: Further Thoughts on a Methodological Classic." Pp. 113–35 in *Émile Durkheim*, edited by R.

Nisbet. Englewood Cliffs, NJ: Prentice-Hall.
Stack, Steven. 1983. "The Effect of Religious Commitment on Suicide: A Cross-National Analysis." *Journal of Health and Social Behavior* 24:362–74.
Stark, Rodney, Daniel P. Doyle, and Jesse Lynn Rushing. 1983. "Beyond Durkheim: Religion and Suicide." *Journal for the Scientific Study of Religion* 11:120–31.
van Poppel, Frans and Jitse P. van Dijk. 1994. "The Development of Cause-of-Death Registration in the Netherlands: 1865–1955." Working Paper No. 30, Indiana University Population Institute for Research and Training, Indiana University, Bloomington, IN.
Veling, Solke H. J. and Fred Sturmans. 1981. "Standaardisering van morbiditeits- en mortaliteitscijfers en indices" (Standardization of Morbidity and Mortality Rates and Indices). *Tijdschrift Sociale Geneeskunde* 59:271–78.