CAMBRIDGE
Cambridge University Press & Assessment
978-0-521-84993-7 — Ethnicity and Causal Mechanisms
Michael Rutter, Marta Tienda
Excerpt
More Information

1

Natural Experiments, Causal Influences, and Policy Development

Michael Rutter

Policy makers, like practitioners and members of the general public, are constantly faced with the need to decide when to take action on the basis of research findings supposedly showing that a particular individual characteristic or environmental circumstance is associated with a markedly increased risk for some negative outcome. Thus, over the years, campaigners have argued for the apparent need to prevent mothers from taking jobs outside the home, or to stop unmarried mothers from having children, or to restrict immigration, or to avoid immunization on the grounds that each of these carried serious risks for the children. But do they? How can we decide which research findings should lead to action and which should not?

In part, that issue involves asking which findings we should believe (is the claimed association real?); in part, it requires consideration of whether the causal inferences are justified; in part, it means questioning whether the proposed risk mechanisms are truly the ones that carry the risk; and, finally, it means considering whether the risks operate generally or only in certain circumstances. These questions constitute *the* major challenge for the whole field of social and behavioral sciences, and my purpose in this chapter is to discuss how they may be tackled. My messages are to caution against uncritical acceptance of claims regarding causal influences but to recognize that good research strategies are available to test causal inferences and to appreciate that these have led to some reasonably solid conclusions.

Hypotheses about possible risk factors that might contribute to the causal mechanisms involved in the origins of some maladaptive, or otherwise undesirable, psychosocial outcome are usually based on some form of group comparison showing that there is a statistically significant association between the putative risk factor and the outcome of interest. Thus, the origin might be evidence that males are more likely than females to engage in antisocial behavior (Moffitt, Caspi, Rutter, & Silva, 2001); or that children experiencing prolonged early group day care outside the family home are

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-0-521-84993-7 — Ethnicity and Causal Mechanisms
Michael Rutter, Marta Tienda
Excerpt
More Information

more likely than those receiving home care to be aggressive (Belsky, 2001); or that schizophrenia is more frequent in those of African-Caribbean background than in Caucasians living in the United Kingdom (Rutter, Pickles, Murray, & Eaves, 2001). As already noted, the four crucial questions in relation to any such evidence are (a) is the association valid?; (b) if valid, does it represent a causal effect?; (c) if there is a causal influence, what element in the experience or circumstance provides the risk and by what mechanism does it operate?; and (d) does the risk operate in all people in all circumstances or is it contingent on either particular individual characteristics or a particular social context? These issues constitute the subject matter of this chapter.

Validity of the Association

Although there are numerous methodological points that have to be considered with respect to the validity of the association between any putative risk factor and the adverse outcome being considered, the two most basic concern representativeness of sampling and comparability of measurement across the groups being contrasted (Moffitt et al., 2001; Rutter, Pickles, et al., 2001a; Rutter & Nikapota, 2002; Rutter, Caspi, & Moffitt, 2003).

Sampling

With respect to sampling, the key need is for representative general population epidemiological samples with a low attrition or nonparticipation rate (Berk, 1983; Sher & Trull, 1996; Thornberry, Bjerregard, & Miles, 1993). Clinic groups or volunteer samples are highly likely to be biased in ways that matter, and individuals who are untraceable or decline to participate in a study tend to be systematically different from those who take part.

Measurement

Comparability of measurement is fundamental for all epidemiological studies. Traditionally, the approach used to be to take some majority group, select the measures that worked best in that group, and then apply those measures to the supposed risk population. However, it has long been obvious that that is potentially biasing. Thus, questions had to be raised as to whether insecure attachment has the same meaning in children experiencing group day care as in those looked after at home by their parents (NICHD Early Child Care Research Network, 1997) or in children suffering severe institutional deprivation (O'Connor et al., 2003) or in children from cultures with very different patterns of parenting (van IJzendoorn & Sagi, 1999). Similarly, there were queries on whether males and females showed their antisocial behavior in the same ways (Moffitt et al., 2001) and on whether cultural and ethnic groups vary in the manner in

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-0-521-84993-7 — Ethnicity and Causal Mechanisms
Michael Rutter, Marta Tienda
Excerpt
More Information

which they expressed their psychopathology (Rutter & Nikapota, 2002). When the measurement of adverse outcomes is dependent on police practice and judicial processing (as is the case with crime statistics – Rutter, Giller, & Hagell, 1998; Williams, Ayers, Abbott, Hawkins, & Catalano, 1996) or on psychiatric diagnosis (see, for example, Hickling, McKenzie, Mullen, & Murray, 1999, regarding schizophrenia in ethnic minorities) it is also necessary to determine whether these procedures operate in the same way across the groups to be compared. In each instance, what is required is systematic validity testing of the measures in *each* of the groups to be studied, and use of the same set of measures across groups, with the set inclusive of what is optimal with respect to sensitivity and specificity for each group.

Statistical Analyses

Given appropriate sampling and measurement, the further need is to undertake suitable statistical analyses. In that connection, the two most basic hazards are (a) the danger of false positives if a large number of possible risk factors are studied in the style of an unfocussed fishing expedition; and (b) the error of concluding that if an association is statistically significant in one group and not in a second group, there is a significant difference among the groups in the association found (see Cohen, Cohen, & Brook, 1995). It does not. The most fundamental point, however, is that (regardless of the level of statistical significance) the only true test in science of the validity of a finding is independent replication of the result in a separate sample by a different group of researchers. Until that happens, policy makers and practitioners should be hesitant about accepting any finding as valid.

Noncausal Alternatives

Before discussing the range of research strategies available to test causal inferences, we need to consider the alternatives to causation once it has been established that there is a replicated valid association to explain. Five main possibilities have to be considered: (a) that the association reflects some form of social selection; (b) that the causal arrow runs in the reverse direction; (c) that there is a causal effect but it is genetically, rather than environmentally, mediated; (d) that it is due to some third variable with which the putative risk factor happens to be associated; and (e) that the risk element has been mis-specified.

Social Selection

The underlying point with respect to social selection is that environments are not randomly distributed (Rutter, Champion, Quinton, Maughan, & Pickles, 1995). For example, being born to a teenage parent is well

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-0-521-84993-7 — Ethnicity and Causal Mechanisms
Michael Rutter, Marta Tienda
Excerpt
More Information

established as an important risk factor for children's psychological disturbance (Moffitt & the E-Risk Study Team, 2002), but it is known that young people who become parents as adolescents are very likely to have shown disturbed behavior or low educational attainments themselves, and it is necessary to ask whether the risks derive from being reared by a teenage parent or from the genetic and environmental risks (for the offspring) associated with the types of teenager who become a parent at an unusually early age. It is evident that similar questions need to be asked with respect to the effects on children of parental divorce, or being brought up by a single parent, or indeed with respect to almost all aspects of child rearing.

In not quite so obvious a fashion, it also applied to studies of ethnicity. Thus, immigrants may represent an atypical sample of the inhabitants of the country from which they have come (Odegaard, 1932), and the operation of housing and job discrimination may mean that ethnic minority families have an increased likelihood of social disadvantage that reflects the response of the host culture to immigrants or ethnic minorities, rather than anything about the immigrants or ethnic minorities themselves.

Person Effects on the Environment

When the putative risk factor concerns any kind of socialization experience, it is always necessary to consider whether the association between that experience and some adverse outcome represents the causal effect of socialization on the child's functioning, or whether, instead, it is due to the child's effect on his/her social environment (Bell, 1968; Bell & Chapman, 1986). There is good evidence that how children behave influences the reactions of other people to them and thereby shapes their environment. Of course, too, to an important extent, children can select the environments they enter. This is most obviously the case with respect to their choice of peer groups (Kandel, 1978; Rowe, Woulbroun, & Gulley, 1994), but the point applies more broadly. This alternative explanation does not apply directly in the case of risks supposedly associated with an individual characteristic such as gender or ethnicity, but it is certainly relevant with respect to many of possible mediating mechanisms associated with the individual characteristic.

Genetic Mediation

The next alternative in relation to any socialization experience is that the risk is mediated genetically rather than environmentally (Plomin, 1994, 1995). This possibility arises because genes affect individual variation in all forms of behavior. This means that any experience that can be influenced by how people behave involves the possibility that the risks are (at least in part) genetically, rather than environmentally, mediated. This applies, for example, to any aspect of parenting, to divorce or single parenthood, and to many types of life stress. The mere fact that a variable is conceptualized

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-0-521-84993-7 — Ethnicity and Causal Mechanisms
Michael Rutter, Marta Tienda
Excerpt
More Information

as "environmental" does not necessarily mean that the associated risks are environmentally mediated.

The same consideration, but the other way round, applies to individual characteristics. Thus, whether a person is male or female is determined genetically, but that does not necessarily mean that any associated risks for psychopathology involve a proximal risk process that is genetically mediated (see Rutter, Caspi & Moffitt, 2003, 2004). Even more so, the same applies to ethnicity. Ethnicity is a complex concept that may be based on religion, history, or geography rather than biology (Rutter & Nikapota, 2002; see also Chapter 3 in this volume). Nevertheless, some identifying ethnic features (such as skin pigmentation) are genetically determined, or at least strongly genetically influenced. But that certainly does not mean that any risks associated with skin color are genetically mediated. Thus, racial discrimination concerns an environmental influence from other people, and not a genetic effect within the individual. The need, as always, is to avoid inferring either genetic or environmental mediation, but instead to use research strategies to test which it is.

Third-Variable Effects

An ever-present consideration in any study of risk and protective factors is whether the demonstrated association in reality reflects some third variable with which the risk factor happens to be associated. Thus, for example, with respect to ethnicity or immigrant status it is essential to consider whether any association might have arisen because the immigrant group tends to be much younger than the population as a whole, or because the ethnic minority sample includes a disproportionate number of individuals without work or living in poverty or in poor-quality housing (Wilson, 1987). In other words, is the association between ethnicity and some adverse outcome really due to a risk effect deriving from one of these other variables? The need in all cases is to consider what these third-variable effects might be, and then to undertake studies of populations that differ in their associations with the other variables. Internal analyses of a single sample can perform much the same task but, for a variety of reasons, they are less satisfactory (see Rutter, Pickles, et al., 2001). Causal inferences become convincing only when it has been shown that the associations are maintained across a diverse range of samples and circumstances.

Mis-specification of the Risk

A closely related possibility is that the risk factor may have been misspecified. For example, with respect to the notion that "broken homes" caused an increased risk of crime, depression, and other forms of psychopathology, it was necessary to undertake epidemiological studies to determine whether the risk was due to parental loss, or to the family

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-0-521-84993-7 — Ethnicity and Causal Mechanisms
Michael Rutter, Marta Tienda
Excerpt
More Information

discord that led to the breakup of the marriage, or to the adverse effects of the breakup on the parenting provided to the children (see Fergusson, Horwood, & Lynskey, 1992; Harris, Brown, & Bifulco, 1986; Rutter, 1971). In these circumstances, it may often be helpful to conceptualize and specify the situations in which there should not be a risk effect if the risk has been correctly specified (Rutter, 1974).

In the case of immigrant status or ethnicity, the possible mis-specifications include features such as preimmigration experiences, religion, experience of racial discrimination, educational/occupational level, family structure, and economic circumstances – to mention just a few possibilities. It should be noted, however, that such mis-specification does not mean that ethnicity is unimportant; rather it points to the need to break down ethnicity according to the differing meanings and different facets (see Chapter 3 of this volume).

Overview: Individual Characteristic Risks and Risk Alternatives

With respect to the risks associated with relatively fixed individual characteristics, such as gender or ethnicity, the alternative of person effects on the environment is nonoperative, but otherwise the alternatives apply. The main difference from the testing of variable individual characteristics (such as intelligence or personality features or pubertal status), or from the testing of environmental risks, stems from the supposedly fixed nature of the characteristic. That has two key implications. First, it is not possible to use the test of a "dose-response" relationship when the risk factor is categorical and fixed. Ordinarily, unless there is a reason to suppose a threshold effect, if there is a true causal effect, it may be expected that the greater strength of a risk factor, the greater the effect on the adverse (or beneficial) outcome. That cannot apply to a fixed categorical feature. However, it needs to be emphasized that the apparently fixed nature of a characteristic is entirely dependent on the risk or protective mechanism that is operative. Thus, in the case of sex, clearly chromosomal sex is fixed and, short of major surgery, so is genital sex. By contrast, sex hormone levels are not fixed and are, in any case, dimensional rather than categorical. This is even more the case with respect to sex roles and societal expectations. Exactly the same applies to ethnicity. Skin pigmentation is fixed and so are the genetic aspects of race (see Chapter 3 of this volume). On the other hand, personal ethnic identification is not fixed and neither is racial discrimination or societal constructions regarding the meaning of ethnicity.

Second, it makes no sense to conceptualize the risk effects of fixed characteristics as operating as a direct proximal risk mechanism. Instead, it is necessary to consider different levels of risk mechanism (Rutter, Caspi, & Moffitt, 2003). Thus, in the case of sex (gender), the basic distal starting point has to be the genetic determination of the biological sex as male or female, because it is that that defines the fixed risk characteristic. The second

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-0-521-84993-7 — Ethnicity and Causal Mechanisms
Michael Rutter, Marta Tienda
Excerpt
More Information

level comprises the varied consequences of being male or female. These consequences are quite diverse – spanning prenatal hormonal effects; hormonal changes in later life; biological effects on physical vulnerability and life expectancy; biologically determined, sex-limited, experiences such as childbirth; and a wide range of culturally influenced experiences that differ between the two sexes (such as the nature of peer groups, living with a male partner, sexual discrimination, and the likelihood of being sexually abused or suffering a head injury). These second-level consequences get one closer to the actual process that leads to psychopathology (or whatever outcome is being considered), but are unlikely to constitute the direct proximal causal risk mechanism. That requires some third-level process that arises out of the second-level consequences. Again, the possibilities are many and various. Thus, with respect to psychopathology, they span personality features such as neuroticism or sensation-seeking, cognitive sets or styles such as a bias towards the attribution of hostile intent or self-blame, or a high susceptibility to certain psychosocial stressors.

It is obvious that parallel considerations will apply to risk or protective features described in terms of ethnicity or immigrant status. The distal starting point will, of course, vary according to the particular concept – be it skin color, religion, geography, or history. Genetic influences will be major for some aspects of ethnicity but much less so for others. Similarly, the second-level consequences will vary according to the concept, but it is necessary to appreciate the diversity of the possibilities. Thus, there are genetic liabilities associated with some ethnic groups (but different from the genes implicated in the definition of ethnicity). For example, there is the genetic propensity to develop an unpleasant flushing response after the ingestion of alcohol that occurs in about a quarter of Japanese individuals but not Caucasians (Ball & Collier, 2002; Heath et al., 2003). The abnormal response derives from a single gene mutation that leads to an inactive enzyme; its psychopathological importance lies in the considerable protective effect against alcoholism that it provides. It is a second-level consequence and not a first-level one because it does not define Japanese ethnicity and because it is present in only some Japanese people. It is a second, not third proximal, level feature because the flushing response is not itself directly involved in the causal process. Rather, it is the affective correlates that are closer to the key mechanisms. The apparently lesser risk of Alzheimer's disease associated with the apoE4 gene in those of African racial background constitutes another example (Hendrie, Hall, et al., 1995; Hendrie, Osuntokun, et al., 1995; Rubinsztein, 1995), as does the sickle-cell trait and its protective effect against malaria (Davies & Brozovic, 1989; Weatherall & Clegg, 2001).

Of course, the genetic consequences of ethnicity constitute but one possibility. In addition, and often of greater importance, are the responses of other people, as evident in racial or religious discriminations, and life-style

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-0-521-84993-7 — Ethnicity and Causal Mechanisms
Michael Rutter, Marta Tienda
Excerpt
More Information

effects, such as those that may involve constraints on females. None of these directly cause psychopathology or even individual differences in psychological traits, but they may make them more likely because of the connections with proximal risk factors. The proximal risk factors are likely to be similar to those found in other groups, so that the key question is why and how they are linked with ethnicity.

Unfortunately, genetic considerations in relation to ethnicity bring forth all sorts of prejudicial responses in many people. In view of the historical abuses associated with eugenics (see Devlin, Fienberg, Resnick, & Roeder, 1997) it is understandable that these attitudes exist. Nevertheless, it is crucial to seek to get the balance right (see Chapter 3 of this volume). In that connection, we need to note that, biologically speaking, "races" are not categorically distinct, and that the genetic similarities among ethnic groups far outweigh the differences. Even so, in particular instances, as the examples given illustrate, the differences may be crucially important (see Risch et al., 2002). We all recognize the severe dangers of inferring a genetic basis for psychological differences among ethnic groups on the evidence of genetic influences on individual differences within ethnic groups (Tizard, 1975). On the other hand, it is quite possible that genetic factors may play a role in some differences among ethnic groups with respect to biologically influenced traits.

For example, just conceivably, that possibility might apply to our findings on height in London children of African-Caribbean origin some 30 years ago (Yule, Berger, Rutter, & Yule, 1975). The children were some 4 cm taller than their Caucasian peers of the same age but, within the children of West Indian parentage, those born in the United Kingdom were some 2 cm taller than those born in the West Indies. The latter finding was probably a function of nutritional differences, but the former might reflect genetic influences on either height or rate of physical maturation. We did not follow up the finding because it was not the focus of our research, but the point is that the combination of within- and between-group differences may suggest possible modes of mediation worth investigating further.

Research Strategies to Test Possible Causal Mechanisms

It has been argued that there are five essential design features (in addition to multiple methods of measurement, the use of longitudinal data, and good statistical methods) that characterize the research strategies needed to put causal hypotheses on mediating mechanisms to the test (Rutter, Pickles, et al., 2001). They are, first, the selection of samples that serve to pull apart variables that ordinarily go together; adoptee strategies, twin designs, natural experiments of different kinds, migration designs, time-series analyses and intervention experiments all serve to do this. Because of the importance of interaction effects, as discussed hereafter, there is

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-0-521-84993-7 — Ethnicity and Causal Mechanisms
Michael Rutter, Marta Tienda
Excerpt
More Information

particular value in designs that can simultaneously "pull apart" and "put together" risk and protective variables. Second, it is necessary to consider the processes that lead to risk exposure. The main problem in causal inferences is the nonrandom assignment of risks, rather than the operation of multiple causes. Third, it is vital to compare and contrast alternative causal mechanisms, rather than test just one favored possibility. Fourth, it is crucial to identify the key assumptions in the chosen design and test whether these assumptions are actually met. Finally, it is helpful to combine consideration of causal influences with respect to both individual differences in liability and group differences in rates or level of the outcome being studied. In the remainder of this section of the chapter, the prime focus, however, will be on the testing of causal inferences on the mechanisms that might be operative in group differences in the level of some psychological trait or disorder. The examples chosen all concern "negative" outcomes of one sort or another, because those are what have been the main subject of research. However, it is crucial to note that ethnic variations concern positive, as well as negative, outcomes (see Chapter 3 of this volume). Indeed, maximum research leverage is obtained by considering both together. The same principles apply. Six different approaches will be used to illustrate the range of the main considerations that need to guide the choice and use of research designs.

Sex Difference in Antisocial Behavior

The first example concerns the use of the Dunedin longitudinal study to investigate the mechanisms that might mediate the well-established tendency for males to be more likely to engage in antisocial behavior (Moffitt et al., 2001). In this case it was necessary to start by checking whether the difference might be an artifact of the two sexes showing their antisocial behavior in different ways or of the need to use a different threshold in males and females. Predictive validity in relation to adult functioning constituted the main test. The findings showed that, although there were some interesting differences in pattern, comparabilities far outweighed differences. But, it was also found that the sex difference largely applied to lifecourse-persistent antisocial behavior, was much less evident for such behavior when it was largely confined to the adolescent years, and was least evident for domestic violence. The focus, therefore, needed to be particularly on the marked male excess for antisocial behavior beginning early in childhood and continuing into adult life. The next research questions concerned the possibilities that the risk factors differed in males and females, that males experienced more (or more severe) risk factors, or that they were more susceptible to the same stressors. In brief, it was found that the key difference was that males were more likely to show early-onset hyperactivity, cognitive impairment, and temperamental difficulty; these were risk factors in both sexes but they were more likely to be

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-0-521-84993-7 — Ethnicity and Causal Mechanisms
Michael Rutter, Marta Tienda
Excerpt
More Information

experienced by boys. When these variables were introduced into a causal model, they eliminated most (but not all) of the sex difference. The research focus now needs to shift to the question of why these risks more often occur in males. In addition, the findings emphasize the likely importance of these risk factors for antisocial behavior more generally in both sexes.

Language Impairment in Twins

The next example concerns the investigation of the causes of the, on average, impaired language development in twins as compared with singletons (Rutter, Thorpe, Greenwood, Northstone, & Golding, 2003; Thorpe, Rutter, & Greenwood, 2003). Three main alternative explanations had to be considered: (a) that the difference was a function of the higher rate of obstetric and perinatal complications in twins; (b) that it was due to some risk factor (such as the transfusion syndrome or an overcrowded womb) that was specific to twins; or (c) that it was a consequence of some altered pattern of family interaction brought about by having to deal with two babies at roughly the same developmental level at the same time. In this case, the comparability issue particularly concerned the fact that twins tend to be born biologically less mature than singletons. This meant that the language outcome had to be adjusted to be in line with the children's age since conception rather than since birth. Also, obstetric risk analysis had to recognize that the optimum gestation period for twin is 37 weeks, rather than 40 weeks as for singletons. Accordingly, it was necessary to standardize within groups in order to examine the effects of unusually short or long gestation.

Having dealt with these measurement issues, four requirements were set for the criteria for a valid inference on causation (with respect to the twin–singleton difference in language): (a) the putative risk variables had to differ significantly in frequency or severity between twins and singletons; (b) the variables had to be significantly associated with language outcome at 3 years within both the twin and singleton samples; (c) this association had to be maintained after taking account of the children's language level at 20 months (this requirement that language progress constitute the outcome was necessary to rule out the possibility that the risk variables were brought about by the language impairment, rather than the other way round); and (d) when introduced into a causal model, the risk variables that met the first three criteria obliterated (or greatly reduced) the twin–singleton difference in language performance.

As it turned out, the only variable that met these four criteria were the indices of mother–child interaction and communication. The finding not only accounted for the twin–singleton difference but also indicated that variations in socialization experiences within the normal range (and not just at an abnormal extreme) could affect psychological outcomes. As with the sex difference in antisocial behavior example, the findings on the group

© in this web service Cambridge University Press & Assessment www.cambridge.org