Development and Psychopathology, **15** (2003), 837–852
Copyright © 2003 Cambridge University Press
Printed in the United States of America
DOI: 10.1017.S0954579403000403

Natural experiments to study the effects of early experience: Progress and limitations

THOMAS G. O'CONNOR
Institute of Psychiatry, London; and University of Rochester

Abstract
Questions concerning the long-term effects of early experience on psychological development in humans continue to stir up controversy. There are, at present, great diversity and contradictions among extant conceptual models and inadequate data to resolve these differences. Limited opportunities to use experimental approaches and the inherent restrictions of observational approaches in research on humans implies that alternative research designs are needed. This article examines the use of "natural experiments" as a method for testing hypotheses concerning the effects of early experience on psychological development. The benefits and limitations of using natural experiments are reviewed and several case illustrations are considered.

The limits of experimental and naturalistic studies to test developmental hypotheses have been appreciated for some time (e.g., Bronfenbrenner, 1979). "Experiments in nature"—naturally arising conditions in which there is a possibility of separating otherwise confounded processes or opportunities to examine processes that for ethical or practical reasons would not have been possible—offer particular methodological advantages over more conventional experimental and naturalistic approaches. The current paper examines the use of "natural experiments" to study the effects of early experience on psychological development in humans. After briefly discussing why natural experiments are needed in this area, the review considers several examples of naturalistic studies, experimental investigations, and natural experiments of the effects of early deprivation on human development. A concluding section considers strengths and weaknesses of these designs and outlines how experiments in nature may be used to advance research into the study of early experiences.

Address correspondence and reprint requests to: Thomas O'Connor, Department of Psychiatry, University of Rochester Medical Center, 300 Crittenden Blvd, Rochester, NY 14642; E-mail: Tom_OConnor@URMC.Rochester.edu.

The Problem and the Need for Natural Experiments

General criticisms of experimental approaches come in a variety of forms. For example, experimental studies have been questioned because the findings obtained under tightly controlled settings may not generalize to the "real world" (in which the setting is not controlled). There are many examples of such limitations, including, in the intervention area, the distinction between efficacy and effectiveness studies (e.g., Chorpita, Barlow, Albano, & Daleiden, 1998). Whereas the former are designed to test the intervention under controlled settings in which, for example, patients are closely screened and the treatment is highly monitored, the latter is concerned with the effect of the intervention in less ideal, or more real life, settings. Evidence that an intervention can work (i.e., proof of concept) does not necessarily mean that it will work when under more routine conditions. This is one reason why much of the effort in current large-scale intervention/prevention studies is focused on

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

recruiting representative samples (i.e., samples representative of those with the disorder of interest), improving take-up rates for the intervention/prevention, monitoring treatment fidelity, retaining participants through the duration of the study, and disseminating the findings to a wide audience (e.g., Sanders, Turner, & Markie–Dadds, 2002). Thus, including experimental manipulation, such as treatment, as part of the research design does not eliminate concerns over making causal claims.

Experimental research on the long-term effects of early exposure to risk is also hampered by the obvious ethical problems with certain kinds of manipulations. Thus, for example, the kinds of manipulations that might be used to rigorously test hypotheses concerning the role of early experience (i.e., the type routinely used in animal investigations) would not be acceptable. There are manipulations that are widely used in experimental laboratory settings with children, such as the use of brief and modest manipulations to the child's environment in the Strange Situation separation–reunion paradigm (Ainsworth, Blehar, Waters, & Wall, 1978). Data supporting the meaningfulness of behavior in the separation–reunion paradigm are now substantial. In contrast, a cloud of uncertainty hangs over other kinds of brief and modest experimental manipulations that have not been the subject of such intense research activity, and it is likely that such alterations in behavior might not generalize across time or place (Bronfenbrenner, 1979). In other words, experimental studies with children are not able to mimic the kinds of risk factors and risk processes that are central to debates concerning the role of early experience.

Criticisms about the role of naturalistic studies to examine the role of early experience are altogether different, but no less severe. The core limitation of naturalistic or observational studies to examine the specific role of early experience is the stability of exposure to adverse environments. That is, naturalistic studies show that many of the most important and widely studied risks for healthy psychological development, including poverty and harsh/punitive parenting, are persistent. Such stability in the environment may come about for a variety of reasons. In addition, considerable stability is also observed when the focus is on individuals rather than the environment. Stability of individual differences for many of the phenotypes of interest to developmental psychopathology is also substantial; in many cases, such as antisocial behavior, this stability is partly attributable to the individual having an active role in maintaining or promoting risk in the environment (e.g., Caspi & Moffitt, 1995). Whatever the mechanisms involved, the implication is that the stability of exposure to high-risk environments makes it difficult, if not impossible in some circumstances, to make causal claims about the early exposure. In other words, early exposure to risk is usually confounded with persistence and severity of risk exposure. More generally, the persistent quality of the risks for psychopathology means that it is difficult to make claims about *timing* of risk exposure, whether that means exposure early or late in development.

Several examples illustrate the problem. For example, exposure to poverty in early childhood may be associated with poor social, cognitive, and behavioral development in adolescence, but finding such an association falls well short of establishing that it was the *specific* exposure to poverty in early childhood that was the mediating variable. It may be that it is the cumulative effects of poverty in childhood that may be the explanatory variable. (Welfare reform legislation seems to be creating a form of natural experiment to examine changes in public assistance and poverty even if does not attain its intended objectives, Duncan & Brooks–Gunn, 2000.)

Similarly, a secure child–parent attachment in infancy may be associated with positive social adjustment at school, but obtaining this association does not provide evidence that it was the secure attachment relationship in infancy that was essential. In fact, insofar as the early experience hypothesis of attachment has been investigated (e.g., Belsky & Fearon, 2002), it seems that the benefits derive from persistence of secure attachment and sensitive parental care rather than secure attachment specific to the first year. However, the further question that must be asked is what is differ-

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

ent about those parent–child dyads that were initially insecure but then changed to secure and vice versa. These groups may differ in many other ways, and so it remains a distinct possibility that differences between groups on measures of adjustment are not (only) attributable to differences in attachment histories.

Research strategies

Recent reviews of research into developmental psychopathology and the causal role of environmental experiences or exposure (e.g., Hinshaw, 2002; Rutter, Pickles, Murray, & Eaves, 2001) make at least two important methodological/conceptual points. The first is that, despite the difficulty often encountered in research testing causal hypotheses, it is possible to obtain substantial leverage to testing competing causal hypotheses. That is, although it is important not to overstate causal claims in research because of the difficulty in "ruling in" one proposed mechanism over all others, this should not be interpreted to suggest that we are unable to test causal hypotheses. We can identify causal mechanisms with some precision provided that the research design is adequate. The analytic logic of testing causal mechanisms has been discussed (e.g., Kraemer, Stice, Kazdin, Offord, & Kupfer, 2001). What is important is that the analytic approaches to defining mediating, moderating, and other processes are reasonably clear when the data permit the testing of alternative models of how risk variables and processes act together. The problem encountered with most studies of early experience and long-term adjustment is that the available data are not amenable to this sort of approach; in other words, careful adherence to the outlined principles would still not provide adequate leverage for testing alternative models of early risk exposure. This point is taken up further below.

The second point of consensus, which qualifies the first point, is that if leverage to test alternative causal hypotheses is to be obtained, then there needs to be convergence among alternative research designs. Exclusive reliance on one design may be as limited as research that considers only one measure of a particular construct (e.g., a certain questionnaire assessment of parent–child relationship quality or observations of parent–child interaction collected only under free play conditions). So, for example, there is now general appreciation that the use of twin designs provides an especially powerful natural experiment to test hypotheses concerning genetic mediation of behavioral/emotional development and risk processes. However, twin designs are not the only natural experiment to test genetic hypotheses; so there is a need to consider alternative methods, including adoption, parent–offspring, and other approaches. Findings from twin studies need to be confirmed with those from other genetically informative designs, and it is only when findings from complementary designs are compared that it is possible to know if the findings from one approach can be generalized. The burden for investigators is to develop research programs that make use of multiple and complementary research designs. The significance of findings from natural experiments need to be interpreted within this larger context of research methodology.

Studies Testing the Early Exposure Hypothesis

This section examines examples of alternative research designs to test hypotheses concerning the long-term effect of risk exposure.

Naturalistic/observational studies

Much of the data that have been considered in debates concerning the role of early experience are based on naturalistic studies. Examples of the kinds of studies that have been examined are reviewed below.

Parental divorce and outcomes in adulthood. There is now compelling evidence that experiencing parental divorce in childhood is associated with a range of adverse outcomes in adulthood, including depression, risk for divorce and unstable intimate romantic relationships, reduced educational attainment, premarital and teenage pregnancy, and selecting a mate with a history of early adversity (such as parental divorce; Amato, 1996; Kiernan, 1992; Kiernan & Hobcraft, 1997; O'Connor,

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

Thorpe, Dunn, Golding, & the ALSPAC Study Team, 1999; Rodgers, Power, & Hope, 1997).

Following on from these robust results, several studies have attempted to discern what it is about the experience of parental separation in childhood that predicts long-term adjustment difficulties. One particular question along these lines is whether the age of the child at the time of separation moderates its effects into adulthood. That is, is there a timing effect underlying the strong patterns of association between parental separation and adult outcomes? Many conflicting and complementary ideas about the vulnerability of children to parental separation have been expressed. For example, compared to preschool and school-aged children, adolescents may adapt better to transitions because they may draw from an enhanced coping capacity and ability to make sense of their parents' separation. Older children also may be protected because they have greater autonomy and access to extrafamilial relationships that may reduce exposure to family conflict. Alternatively, adolescence may instead be seen as a time of increased vulnerability, especially because they are experiencing other normative transitions, such as puberty and moving to secondary school. Experiencing parental separation, courtship, and repartnering may be especially difficult for adolescents who are negotiating their own opposite-sex intimate relationships. In further support of this position, it could be argued that infants and young children may not process (e.g., remember) the divorce/separation and the associated conflict and hence may not be "affected" to a degree similar to older children who retain memories of the divorce process—so long as there is no repeat of conflict and separation.

Unfortunately, despite many years of investigating children's adjustment to parental separation, the role of child age at separation in long-term adjustment remains unclear (e.g., see Hetherington, 1999); the findings are typically contrary or not significant. The reasons for this are not surprising. Children who experience parental separation in early childhood differ in several important ways from those who experience the same event in late childhood, adolescence, or adulthood in terms of, for example, likelihood of experiencing a remarriage, subsequent separation, and subsequent repartnering (i.e., multiple transitions). Those children who experience a parental separation in middle to late childhood may differ from those who experience the same event in early childhood in terms of duration of exposure to marital conflict. Thus, although it would be possible to compare children in different families who experience parental divorce at different ages, there are likely to be other factors distinguishing these children and families that would confound rigorous ("experimental") comparisons. After all, divorce itself is a complex phenomena that is predicted from a range of individual personality, relationship, and social context variables (Amato, 1996). Differences between the groups of children with early versus later exposure to separation need to be controlled (experimentally or naturally) if progress is to be made in understanding the role of child age in adjusting to parental separation.

One possible solution might be to analyze the effect of child age as a source of within-family differences in children's adjustment to parental separation. That is, rather than assess child age across families, we should assess the prediction of child age within families (i.e., whether siblings of different ages adjust differently to the same parents' separation, e.g., O'Connor, Dunn, Jenkins, Pickering, & Rasbash, 2001). There are other problems with this approach, including the possibility that siblings within the same family have been differentially exposed to marital conflict or other strains that precede separation. Nevertheless, siblings have been used as effective controls, especially in genetic studies (Farmer, Mahmood, Redman, Harris, Sadler, & McGuffin, 2003).

Maternal depression. A second example in which the effects of early exposure have been deciphered from naturalistic studies concerns maternal depression. As with the case of parental separation, there is now a robust connection between maternal depression and a range of social, cognitive, and behavioral/emotional problems in children (Brennan, Hammen, Andersen, Bor, Najman, & Williams, 2000; Essex, Klein, Miech, & Smider, 2001;

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

Murray, Fiori–Cowley, & Hooper, 1996). What is a matter of debate is the extent to which postnatal depression, in particular, may be pathogenic. That is, are there particular effects of maternal depression in the infant's first months of life that have lasting effects or is it the case that postnatal depression effects on children's development, if they persist, are mediated by subsequent risk, including further episodes of maternal depression?

There are good reasons for thinking that postnatal depression may be especially harmful for children's psychological development to the extent that the depression interferes with the normally developing child–parent attachment relationship. To this end, Murray and colleagues (1996) reported that maternal depression at 4 months was associated with less sensitive parental behavior and an increased likelihood of an insecure attachment at 18 months. (It was the early assessment of depression that appeared to play an especially predictive role.) Subsequent assessments of that sample have strengthened the claim that postnatal depression at 4 months is associated with long-standing and persisting behavioral/emotional problems in children (Sinclair & Murray, 1998). On the other hand, other studies of maternal depression in the postnatal period and after have been less successful in detecting an effect of postnatal depression per se or, more generally, a timing effect. For example, Kurstjens and Wolke (2001) found that it was concurrent depression rather than postnatal depression that was associated with children's behavioral/emotional problems and social adjustment at age 6 years.

As with the case of parental separation, research findings concerning the effect of age at exposure or early exposure to maternal depression are inconsistent. Contrasts may be made between children exposed to maternal depression uniquely in the postnatal period and children exposed to maternal depression only outside the postnatal period; however, even if this were possible (such groups would exist in a sufficiently large sample), the concern is that these groups of mothers and children might differ for other reasons and, if so, this would confound direct comparisons between groups.

Child–parent attachment. A final example of how naturalistic/observational studies have been used to make causal claims concerning the long-term effects of the child's early environments is in the area of attachment. There has been a tendency to interpret the attachment data, much of which involves mothers and infants, to suggest that a secure attachment in infancy is directly associated with long-term positive adjustment in children. That may be an overstatement (e.g., see Bruer, 1999). Data from the NICHD study of early child care (Belsky & Fearon, 2002) revealed that there were some children who were rated as demonstrating a secure attachment at 15 months but experienced insensitive care at 24 months (attachment security at 24 months was not included in the above report). The finding that there was some degree of discontinuity in attachment/sensitive parenting is not surprising. Indeed, it was desirable because it allowed the investigators to examine children with different histories of Secure attachment and sensitive care. Analyses indicated that, as expected, children with persistently positive child–parent relationship quality exhibited optimal development. More interesting, perhaps, is the finding that children who exhibited an insecure attachment at 15 months but who had mothers who were rated as sensitive at 24 months were judged to show better overall adjustment (e.g., social competence, problem behavior, school readiness) than children who displayed a Secure attachment at 15 months but who had mothers who were rated as insensitive at 24 months. In other words, quality of child–parent attachment assessed close in time to child adjustment was more strongly associated with outcome than early child–parent attachment. Direct contrasts between the two groups rejects an early sensitive period hypothesis for child–parent attachment. However, if the two discontinuous groups differ on variables other than histories of attachment relationships, then direct contrasts between the two would be confounded and any conclusions regarding early versus later effects of a secure attachment would be undermined.

The three empirical illustrations given above serve as examples of how findings from naturalistic studies have been used to make con-

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

clusions concerning the effects of early exposure to risk. However, these studies are not necessarily positioned to test an early experience hypothesis, and so there are potentially important methodological drawbacks in each. Critically, each of the above studies relies on the existence of a naturally occurring change within the sample. Thus, not all the depressed mothers stay depressed throughout the postpartum period and some develop depression after the postpartum period. Similarly, many studies have now demonstrated that there is considerable movement between Secure and Insecure attachment, even over comparatively short time periods (e.g., Vaughn, Egeland, Sroufe, & Waters, 1979; Vondra, Shaw, Swearingen, Cohen, & Owens, 2001), although there may also be stability across two decades (Waters, Merrick, Treboux, Crowell, & Albersheim, 2000). The above studies use this within-sample variation to contrast different subgroups and assume that these different subgroups are comparable, yet, this is unlikely to be the case. Thus, Vaughn et al. reported that those children whose attachment classifications changed between 12 and 18 months could be differentiated on life events; a somewhat similar observation was made by Belsky and Fearon (2002). In other words, the change that these studies depend on is not random but linked to other levels of stress exposure (not under the control of the investigators) that, in turn, may confound comparisons between groups. What is needed are studies in which the change in risk status is imposed from without, as in the case of a randomized clinical trial or a risk exposure not under the control of and independent of the characteristics of the individuals in the sample.

Experimental studies assessing the role of early exposure to adversity

To date, the most persuasive data concerning the role of early experience on human development using an experimental method are the studies of early intervention. Many studies using experimental leverage have now shown that children who receive early social, academic, nutritional, and social care interventions fare better than comparably matched children who had not received these interventions (e.g., Barnett, 1995; Campbell & Ramey, 1994; Reynolds, 1994). However, the extent to which the benefits derive from, and are specific to, *early* intervention has not been resolved (e.g., Bruer, 1999).

Experimental interventions are especially powerful for testing hypotheses concerning the effects of early experience, and they have a number of advantages over the naturalistic studies cited above. Nonetheless, there are surprisingly few examples. Interventions have instead been devised to test hypotheses such as whether early interventions work, to what extent multimodel interventions are needed, and how persistent are the effects of early intervention (e.g., Campbell, Pungello, Johnson, Burchinal, & Ramey, 2001). Admittedly, the question of whether (and for whom) an intervention is effective needs to be addressed before turning to more subtle and complex questions concerning timing. Moreover, attempts to show that an intervention is effective have proven to be difficult enough in most cases. Accordingly, a move toward using intervention/prevention studies to examine timing will require input and impetus from a broader research perspective that considers how the use of interventions, and especially preventive interventions, may tackle questions of basic significance to developmental theories of psychopathology (Cicchetti & Hinshaw, 2002; Dawson, Ashman, & Carver, 2000).

Significantly, one of the obstacles to hurdle in this area is the reluctance to challenge the generally accepted position that "earlier is better." The point is not to undermine clinical experience attesting to some potential benefits of early intervention or reject the possibility that it might be most humane to address problems when they first arise rather than wait for acute disturbance. Instead, as Clarke and O'Callaghan (2003) and Birchwood and Spencer (2001) note for schizophrenia, the aim is to view the earlier is better position as a hypothesis that requires testing. Furthermore, even if there were strong empirical support for intervention in the prodromal stage, the case for early intervention may not be straightforward. There may be, for example, drawbacks of early intervention, where "early" is defined as

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

young children or in terms of the stage of illness. For instance, there is a set of ethical issues that arise when individuals are targeted for treatment despite being in the prodromal rather than disordered state, perhaps especially if medication is involved (e.g., McGlashan, 2001).

Natural Experiments to Study the Effects of Early Experience

Three types of examples of natural experiments to test hypotheses concerning early experiences are reviewed. The first is based on children exposed in utero to high levels of radiation; the second concerns the effects of starvation on long-term psychiatric illness in adulthood; the third involves children rescued from very depriving institutional settings.

In utero exposure to radiation

The Chernobyl nuclear power plant accident in 1986 was an environmental disaster that also had severe health consequences for a large population. Notwithstanding the tragedy of the event, several groups of researchers have sought to capitalize on this natural experiment to examine the effects of high radiation exposure on health outcomes. Of particular significance is the fact that the exposure to radiation could be precisely timed and children at different stages of development could be studied to assess a possible moderating effect of age on radiation exposure; in addition, there were no or minimal selection effects (i.e., characteristics of the individuals affected were not systematically associated with exposure levels). From the scientific perspective, these are the "experimental" features of this natural experiment.

Two studies, in particular, examined radiation exposure in utero and cognitive and behavioral development in middle childhood. One study (Nyagu, Loganovsky, & Loganovskaja, 1998) examined children born between April 1986 and February 1987. Two groups of children were selected according to geographic location (i.e., level of exposure), and there was a further group of children born to mothers who had been evacuated. Children in these groups were matched for age and gender; data were taken from parent and teacher reports and clinical data of thyroid function. Nyagu et al. (1998) found an increase in mental retardation, behavioral and emotional problems, and problems in thyroid function in the children exposed to radiation in utero. Based on their findings, the authors propose that brain abnormalities, perhaps in the limbic–reticular structures in the left hemisphere, account for the pattern of findings; they also suggest that there may be a sensitive period of the developing brain to exposure between 8 and 25 weeks gestation. A complementary study by Igumnov and Drozdovitch (2000) compared two groups of children (6–7 and 10–12 years) who had been exposed in the prenatal period at the time of the Chernobyl disaster to children from less contaminated areas; they found somewhat less marked differences than Nyagu et al.

Numerous studies have suggested that exposure to pathogens of many kinds in utero (e.g., various substances such as cocaine, various viruses such as rubella) is associated with compromised brain development and leads to long-term cognitive and behavioral abnormalities (e.g., Smotherman & Robinson, 1990). However, in most cases, exposure to particular pathogens cannot be precisely timed because, for example, drug-addicted pregnant women will expose the fetus to drugs at more than a single time point (e.g., Molitor, Mayes, & Ward, 2003). The significance of the above studies following the Chernobyl disaster are important in demonstrating how severe radiation may impair brain development, but they are more important in suggesting that there may be periods of brain development that may be especially vulnerable to the effects of radiation.

Prenatal famine and stress and risk of affective disorder in adulthood

Data from a Dutch birth cohort were used to examine the association between prenatal famine and risk for affective disorder (e.g., Brown, van Os, Driessens, Hoek, & Susser, 2000). The situation that created a natural experiment

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

was the Dutch hunger winter (1944–1945) following the Nazi blockade during World War II. Brown et al. found that the rate of affective disorder (unipolar and bipolar depression) in adulthood requiring hospitalization was significantly higher among those whose mothers would have experienced starvation during the second and third trimesters of pregnancy. In addition to showing a link between prenatal famine and risk for affective disorder in adulthood, the study design also allowed for a test of a timing hypothesis.

From an experimental perspective, the key feature of the study was the ability to contrast the rates of disorder among individuals whose mothers were or were not pregnant during the famine. Furthermore, in this case, a severe risk factor (starvation) that is ordinarily stable was instead time limited and occurred for a clearly definable period.

Watson and colleagues (Watson, Mednick, Huttunen, & Wang, 1999) reported complementary findings from a study of children exposed to severe earthquake in Tangshan, China. These authors found that males exposed to the earthquake during the second trimester exhibited higher than expected rates of depression in adulthood.

Experimental leverage to precisely time the beginning (and to a lesser extent the end) of exposure to risk is a particular advantage of natural experiments reported by Watson et al. (1999) and Brown et al. (2000) and is rarely observed or obtainable in naturalistic studies. In fact, several studies of humans (e.g., Glover & O'Connor, 2002) have sought to link maternal stress or anxiety in pregnancy with disturbance in the child, a link that has been established for some time in animals (Schneider & Moore, 2000; Thompson, 1957). However, evidence of a clear timing effect in the human investigations is unclear, largely because of the stability of anxiety and stress. Statistical analyses to covary the effects of stress in early versus late pregnancy are of limited use and are unable to satisfactorily account for other differences that may exist between women with different histories of stress/anxiety in the transition from prenatal to postnatal development.

Studies of children adopted following severe deprivation

A potentially rewarding approach to the study of early experience on long-term development concerns children whose caregiving experience undergoes a radical change. This does not ordinarily occur, but there is an exception to this rule, namely, children who are removed from very abusive homes and placed into nonabusive or therapeutic settings. Children in the adoption and foster care system have garnered increased attention in recent years not only because of the growing realization of the very poor outcomes of these children but also because this situation may provide leverage to study the effects of early adverse parenting that is not confounded with persistently poor parenting (Dozier, Stovall, Albus, & Bates, 2001; Milan & Pinderhughes, 2000; Rushton, Treseder, & Quinton, 1995). One striking example of this kind of study are investigations of children adopted into low-risk family environments after a period of institutional care. One such example is addressed below.

The English and Romanian Adoptee Study. The adoption by U.K. parents of children reared in gross deprivation in Romania was a humanitarian response that was met by questions about the long-term adjustment of these children. The English and Romanian Adoptees (ERA) study was formed to research the numerous basic and more applied issues that the study provided. Of particular significance is that the study offered a natural experiment to examine basic and largely unresolved questions concerning risk and resilience in development and the role of early experience in psychological development. For example, would children who experienced severe early deprivation (and who showed severely compromised development as a result) be able to recover fully after being placed into a low- to normal-risk home? How varied were children in the rate and extent of recovery? What were the distinguishing features of adoptive families in which children showed the most complete recovery?

The sample includes children adopted from Romania and a comparison sample of early-

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

placed adoptees within the United Kingdom. Children in the Romanian sample were drawn from the 324 children adopted into U.K. families between February 1990 and September 1992 who were processed through the Department of Health and/or the Home Office. A stratified random sampling design was used. The final sample included 58 children placed between 0 and 6 months, 59 children placed between 7 and 24 months and a late-placed group of 48 children adopted from Romania and placed between 24 and 42 months. Eighty-one percent of the adoptive parents of Romanian children who were approached agreed to participate. The U.K. sample consisted of 52 children who were placed into adoptive families between 0 and 6 months of age; none of these children experienced deprivation, although there was variation in their backgrounds and reasons for the adoption. Intracountry adoptees were obtained through adoption agencies and social services departments. Available information suggests that approximately 50% of the families who were contacted agreed to participate.

In the case of the ERA study, how successful was the (natural) experimental design? Two sampling considerations might be raised here: first, how representative are the families who adopted from Romania to the general population of families; second, are the families who adopted from Romania comparable to those who adopted within the United Kingdom? Given that the study capitalized on a natural rather than planned experiment, there were some features of the sampling that could not be controlled. For example, adoptive families tend not to resemble the population as a whole. In the ERA study, as with virtually all adoptive studies (a notable exception being Duyme, Dumaret, & Tomkiewicz, 1999), families who adopted children from Romania were, like those parents who adopted from within the United Kingdom, generally middle-class and slightly better educated than the general U.K. population; families of Romanian and U.K. adoptees did not differ from one another, however. As a consequence, we were not able to examine the development of children adopted into high-risk families.

Although families adopting from Romania and within the United Kingdom did not differ with respect to socioeconomic setting, there were other differences between the families of Romanian and U.K. adoptees. Differences that did exist were explained by the fact that some of the families adopting from Romania did not satisfy U.K. conventions applied to baby adoptions (but these did pertain to the families adopting within the United Kingdom). For example, because age differences between adoptive parents and children influence placement decisions, parents of U.K. adoptees (all of whom adopted infants) are significantly younger than parents of Romanian adoptees, by about 2 years for mothers and fathers (at the time of the age 6 assessments, the mean ages for fathers were 42 and 44 years for U.K. and Romanian adoptees, respectively; for mothers, mean ages were 40 and 42 years, respectively). In addition, a greater proportion of families of Romanian adoptees had biological children prior to the adoption (33 vs. 2%), and they were less likely to have adopted previous to the current adoption (4 vs. 48% in the Romanian and U.K. families, respectively). These differences highlight the preference given to prospective intracountry adoptive parents who do not have children (i.e., among those who pursue baby adoptions) and the fact that infertility was not the primary motive among many parents of Romanian adoptees.

Thus, in some ways we were not able to—and, more important, would not have been able to—match families adopting from Romania to the general population or even to those adopting within the United Kingdom. The question, then, is whether these differences in samples are associated with child outcome. The answer is no (in addition to providing some scientific leverage, this finding may have policy lessons concerning what criteria may be relevant for screening prospective adoptive parents).

Another form of potential bias would be introduced if parents who adopted infants from Romania differed from those who adopted older children. This is a form of selective placement, something that is a major concern of adoption studies. As far as we can tell, the

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

answer here is no as well. For example, adoptive parent education was unassociated with either the age at which children were adopted or their progress in the adoptive home (Rutter, O'Connor, & the ERA Research Team, in press). Generally, among the families with Romanian adoptees, no association between family characteristics and child's age at entry into the United Kingdom was found. Nor was there evidence that children who entered the United Kingdom at a relatively young age were any different from those who entered later, in terms of age placed into an institution (most were institutionalized in the first weeks of life) or physical condition at entry (most children were severely malnourished; O'Connor et al., 2000; Rutter & the ERA Research Team, 1998).

In summary, there is considerable evidence that, to the extent that there are differences between U.K. and Romanian adoptees or differences among Romanian adoptees according to the age at which they were adopted, these differences could not be accounted for by differences in the families into which the children were adopted.

One of the commonly asked questions about the study's findings involves their generalizability. That is, as interesting as the study may be, do the findings hold significance and relevance beyond those children who experienced early institutional deprivation and who were adopted internationally? Furthermore, in light of the fact that the conditions in Romania were widely reported to be more severe than those in countries from which international adoption is somewhat common, might it be that the relevance of the findings are restricted to the most severely deprived groups of children and not the growing numbers of children who are adopted internationally from less deprived (but still institutional) settings (e.g., Gunnar, Bruce, & Grotevant, 2000)? A key task was to establish if the findings from the ERA study hold lessons for children who experienced other kinds of early adversity, such children placed in the foster care system or those who experienced severe abuse and neglect (e.g., Minty, 1999; Shealy, 1995).

There are several reasons why the generalizability of the study's findings might be questioned. For example, at the time the study began more than a decade ago, international adoption was comparatively uncommon in the United Kingdom, at least when compared to other European countries (Haworth, 2001). That remains the case. If the needs of parents who adopt internationally are different from those in domestic adoption, then these parents may have particular difficulty in getting their needs understood and addressed. In addition, when the study began, the Romanian situation was thought to be exceptional in the extent to which children experienced deprivation. For example, experience with international adoption in China, Russia, and elsewhere suggests that the children from these settings experienced less severe deprivation (although by no means a normal rearing environment; Gunnar et al., 2000).

Notwithstanding the obvious differences between the children adopted from Romania, results to date suggest several grounds for generalizability. Four sets of findings deserve special attention in that regard. First, as many previous studies of institutionalized children reported (even those of children less severely deprived than those in institutions in Romania, e.g., Tizard & Rees, 1975), we, too, found that many of the children exhibited a form of disturbance in social relationships akin to what is now referred to as an "attachment disorder," rather than more typical forms of insecure attachment (O'Connor et al., 2003). The similarities between the social impairments observed in the children adopted from institutions in Romania and the reports of other investigators were striking, especially considering how different these behaviors were compared to children in other high-risk settings (O'Connor et al., 2003). Second, another form of behavioral disturbance commonly found in the ERA study is inattention/overactivity, a disturbance that has been widely observed among other samples of institutionalized children (Kreppner, O'Connor, Rutter, & the ERA Study Team, 2001). Third, despite the evidence of persisting deficits in some areas, recovery in physical development (e.g., weight, head circumference) was very strong and nearly complete (O'Connor et al., 2000).

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

A similar pattern of complete to near-complete catch-up has been reported in other samples of children adopted into families following institutional deprivation (Lien, Meyer, & Winick, 1977). Fourth, there was a strong, dose–response association between duration of exposure to deprivation and cognitive ability (O'Connor et al., 2000). Indeed, the dose–response pattern is observed for virtually all of the outcomes for which there is an effect of institutionalization.

Each of the four findings reported above suggest that there is probably little difficulty in generalizing the findings from the ERA study to other samples of institutionalized children. In addition, at least in terms of behavioral and social relationship outcomes (but not, perhaps, to cognitive and physical health), this generalization extends to institutional settings characterized by less severe deprivation and, in some respects, even adequate care, such as those described in studies by Tizard and Rees (1975) and by Vorria et al. (in press). In these two studies the children lived in institutions in which adequate nutrition was available and there were opportunities for social interactions and play. The main deficit in these institutions was the lack of consistent care.

What was less clear is how the findings from the ERA study would generalize to larger populations of children who experienced early adversity; more broadly, what lessons do the findings contain for understanding normal development?

The question of generalization has some subtle features that we are still unraveling. For example, there are some differences between certain problems reported in children in the ERA study and children in other high-risk settings. For example, children who exhibited elevated rates of inattention/overactivity at age 6 years did not also exhibit co-occurring conduct problems, a finding that contrasts with what is usually found in high-risk samples (Kreppner et al., 2001). We do not yet know if the inattention/overactivity phenotype found in children who experienced early institutional deprivation differs from that found in ordinary child psychiatry clinics; there is anecdotal evidence that it might differ. If that is so, then there may be different mechanisms accounting for inattention/overactivity in noninstitutional and institutional samples and, equally important, these forms of inattention/overactivity may require different treatments. That possibility is currently being investigated.

A second outcome that raises questions about generalizability from the ERA study is attachment. Disturbances in attachment relationships in children who experienced institutional deprivation have been discussed for many years (e.g., Zeanah, 1996). Attachment disturbances in other high-risk samples (e.g., children of maltreating parents) have also been widely discussed (Cicchetti, Toth, & Lynch, 1995). It is interesting, however, that the kinds of attachment disturbances found in institutional samples and other more "typical" high-risk samples may be different in kind. In fact, in the ERA study, there was no increased risk for the kinds of Insecure attachment that are found in other high-risk samples (Avoidant, Ambivalent/Dependent, Disorganized, Controlling); instead, there was an elevated risk of what has been referred to (by *DSM-IV*) as a disinhibited form of attachment disorder (O'Connor et al., 2003). On the other hand, despite the number of studies assessing attachment insecurity in high-risk samples, the mention of a disinhibited form of attachment disorder is remarkably rare. The implication from the analysis of the attachment data is that the form of disturbance may be particular to the kind of risk to which the children were exposed (i.e., institutional care) and may not generalize to other high-risk samples.

What is clear at present is that concerns about generalizability must be geared, not only to populations (e.g., deciding whether data from institutional adversity is comparable to noninstitutional deprivation such as severe maltreatment) but also to specific clinical problems (findings may generalize to other high-risk samples in the case of inattention/overactivity but not that of attachment disturbances). The implication of findings reported so far is that the ERA study may hold lessons for some areas of developmental psychopathology but not others. To date, we do not yet know the limits of generalizability; this is a topic we are currently pursuing. Understand-

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

ing the limits of the study's generalizability is the key to understanding the essential lessons the study can tell us about normal and atypical development.

As a natural experiment, the ERA study has substantial methodological advantages for researching the role of early adversity in long-term development. In particular, the study is noteworthy because it has followed a representative sample of children whose experience of early severe deprivation was limited to the first months or year of life. This marked discontinuity in risk exposure is what allows us to focus on early risk per se in a way that is uncommon among the many studies following high-risk samples. It is clear that the study holds important lessons for assessing the persistence (or otherwise) of early adversity on long-term social, cognitive, and behavioral/emotional outcomes. The broader question of how the study may shape developmental theory, clinical practice, and policy for young children remains to be settled.

Progress and Limitations of Natural Experiments

The above examples of natural experiments highlight the benefits of this type of design for investigating processes that would have otherwise been difficult or impossible to assess. However, each of the examples provided also exposes different limitations of this design. These are considered in turn.

First, it is evident from the examples considered in this review that natural experiments cannot be constructed to suit all research questions. That is, research using natural experiments is not planned in the ordinary sense. It may well be that some important research questions in developmental psychopathology may not be amendable to natural experiments (i.e., natural experiments may not be available). Nonetheless, there are ample occasions to use many different kinds of naturally occurring experiments at many levels, from social policy to family process. Faced with the obvious difficulties of constructing a study with the necessary experimental leverage, researchers may need to be more opportunistic.

Second, questions about the generalizability of findings from natural experiments cannot always be adequately resolved. This point is illustrated in the findings from the Chernobyl disaster. The results from the Nyagu et al. (1998) and Igumnov and Drozdovitch (2000) studies generally support a connection between radiation and behavioral and cognitive problems and brain abnormalities. What is less clear is whether the findings suggest general lessons concerning the developing brain in utero and some of the possible causes of abnormalities, or alternatively, whether these studies merely show that there are brain abnormalities in response to strong doses of radiation that would not be experienced except in conditions of a severe nuclear accident. If the former were the case, then there would be substantial interest in expanding this area of research and maintaining close connections with, and funding for, these precious samples longitudinally. On the other hand, if the latter were correct, then the findings would be trivial and it would be difficult to justify further resources for follow-up research with these samples.

Understanding how normal and abnormal development may be seen within the same conceptual model is a key feature of developmental psychopathology (Cicchetti & Cohen, 1995); the difficulty of this task is magnified in studies of natural experiments. What seems most likely is that natural experiments (and more specifically the samples in those experiments) have limited but important generalizability. It will be important to define the limits of this generalizability. As shown in the case of the ERA study, this may be a complicated task, and may change as longitudinal data are collected. Natural experiments used to reveal potentially important findings cannot also decipher their generalizability and significance. That requires the appeal to other studies and samples. Thus, for example, understanding the meaning of the attachment disturbance findings from the ERA study requires a comparison of attachment problems in other high-risk samples.

Third, necessarily, natural experiments require replication (insofar as that is possible)

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

and follow-up investigation to test possible mechanisms. Thus, for example, in the case of the natural experiment on prenatal famine capitalized on by Brown et al. (2000), the study was able to establish a connection that appears to be robust. However, what this natural experiment cannot do is test plausible mechanisms accounting for this association. Instead, subsequent clinical and experimental studies are needed to identify the mechanisms mediating the link between prenatal famine and the physiological pathways leading to affective disorder. This is one demonstration of why there will always be a need to bring together a range of research designs. Natural experiments, like epidemiological investigations, provide a stimulus and direction for complementary research testing alternative mechanisms.

Fourth, many kinds of natural experiments may be far less elegant than is desired. A good example concerns studies of children in the adoption and foster care system. Ostensibly, research on children in these settings would allow us to understand the formation of new attachment relationships past infancy, the carrying forward of effects of adverse early (attachment) relationships on subsequent (attachment) relationships, and the factors associated with positive versus negative adjustment to the foster care system. Each of these questions is hugely important, both practically and conceptually. However, in practice, research of this kind is very difficult. For instance, samples of children in the foster care system are enormously heterogeneous in terms of preplacement history, age at removal, time since removal, and biological and social parenting histories. It would be unrealistic to try to estimate the effects of each of these background factors, even in large samples. Moreover, large samples of children are often difficult to secure because of the difficulty in coordinating the many services and individuals involved in the child's care (not to mention the resistance to this kind of research that is sometimes encountered). There is, in addition, a real question concerning the best comparison group or groups. It may be that multiple kinds of comparison groups are needed, as indicated by the work of Tizard and colleagues (e.g., Hodges & Tizard, 1989).

Conclusion

Research into the mechanisms by which early exposure to risk confers long-term disadvantage (where that is the case) requires a collection of research designs and instruments. Given the methodological limits of naturalistic/observational studies and the difficulties in constructing conventional experiments with adequate methodological leverage, research needs to capitalize on naturally occurring experiments. Natural experiments therefore represent an important research tool for developmental psychopathology.

Natural experiments can provide an ability to de-confound alternative and overlapping hypotheses. More specifically, one feature that was especially highlighted in this review is how natural experiments may be used to test alternative hypotheses concerning the timing and/or duration of risk exposure and individual development. Evidence was reviewed from several studies of prenatal and early postnatal stress exposure to suggest that early exposure may have lasting effects. Each study had important limitations and, inevitably, each was concerned with different populations and "naturally occurring" disasters or adversities; nonetheless, collectively they reject the hypothesis that there are no lasting effects of early exposure to risk. Studies that provided the strongest evidence for the early effects of risk on development were not also able to differentiate among several alternative mechanisms to account for this effect, but they did point to the kinds of alternative and follow-up studies that are needed. The implication is that, however important natural experiments may be to test basic questions of early experiences and later adjustment, there remains an important and complementary role for other forms of investigation.

Questions of generalizability were highlighted as a potential limitation of natural experiments, yet one that is often difficult to assess. On the one hand, it may be that natural experiments may have far-reaching applica-

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

tions to inform theories of normal and atypical development across several domains. On the other hand, the application of findings from natural experiments to developmental theory may be very circumscribed and pertain to a limited set of populations and/or a restricted set of outcomes. Findings from the ERA study were highlighted to show the difficulties in assessing the generalizability of findings from natural experiments. At present, we do not have a database adequate for testing the generalizability of natural experiments; this remains an important focus for further research. The integration of findings from several studies that include a diverse array of comparison groups may therefore be essential for placing findings from natural experiments in context and for mapping the generalizability of these findings for developmental theory and applications.

References

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the Strange Situation*. Hillsdale, NJ: Erlbaum.

Amato, P. R. (1996). Explaining the intergenerational transmission of divorce. *Journal of Marriage and the Family, 58,* 628–640.

Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of Children, 5,* 25–50.

Belsky, J., & Fearon, R. M. (2002). Early attachment security, subsequent maternal sensitivity, and later child development: Does continuity in development depend upon continuity of caregiving? *Attachment and Human Development, 4,* 361–387.

Birchwood, M., & Spencer, E. (2001). Early intervention in psychotic relapse. *Clinical Psychology Review, 21,* 1211–1226.

Brennan, P. A., Hammen, C., Andersen, M. J., Bor, W., Najman, J. M., & Williams, G. M. (2000). Chronicity, severity and timing of maternal depressive symptoms: Relationships with child outcomes at age 5. *Developmental Psychology, 36,* 759–766.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Cambridge, MA: Harvard University Press.

Brown, A. S., van Os, J., Driessens, C., Hoek, H. W., & Susser, E. S. (2000). Further evidence of relation between prenatal famine and major affective disorder. *American Journal of Psychiatry, 157,* 190–195.

Bruer, J. T. (1999). *The myth of the first three years*. New York: Free Press.

Campbell, F. A., Pungello, E. P., Johnson, S. M., Burchinal, M., & Ramey, C. T. (2001). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology, 37,* 231–242.

Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children in low income families. *Child Development, 65,* 684–698.

Caspi, A., & Moffitt, T. E. (1995). The continuity of maladaptive behavior: From description to understanding in the study of antisocial behavior. In D. Cicchetti & D. Cohen (Eds.), *Developmental psychopathology* (Vol. 2, pp. 472–511). New York: Wiley.

Chorpita, B. F., Barlow, D. H., Albano, A. M., & Daleiden, E. L. (1998). Methodological strategies in child clinical trials: advancing the efficacy and effectiveness of psychosocial treatments. *Journal of Abnormal Child Psychology, 26,* 7–16.

Cicchetti, D., & Cohen, D. J. (1995). Perspectives in developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology* (pp. 2–20). New York: Wiley.

Cicchetti, D., & Hinshaw, S. P. (2002). Editorial: Prevention and intervention science: Contributions to developmental theory. *Development and Psychopathology, 14,* 667–671.

Cicchetti, D., Toth, S. L., & Lynch, M. (1995). Bowlby's dream comes full circle: The application of attachment theory to risk and psychopathology. In T. H. Ollendick & R. J. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 17, pp. 1–75). New York: Plenum Press.

Clarke, M., & O'Callaghan, E. (2003). Is earlier better? At the beginning of schizophrenia: Timing and opportunities for early intervention. *Psychiatric Clinics of North America, 26,* 65–83.

Dawson, G., Ashman, S. B., & Carver, L. J. (2000). The role of early experience in shaping behavioral and brain development and its implications for social policy. *Development and Psychopathology, 12,* 695–712.

Dozier, M., Stovall, K. C., Albus, K. E., & Bates, B. (2001). Attachment for infants in foster care: The role of caregiver state of mind. *Child Development, 72,* 1467–1477.

Duncan, G. J., & Brooks-Gunn, J. (2000). Family poverty, welfare reform, and child development. *Child Development, 71,* 188–196.

Duyme, M., Dumaret, A. C., & Tomkiewicz, T. C. (1999). How can we boost IQs of "dull children"?: A late adoption study. *Proceedings of the National Academy of Sciences, 96,* 8790–8794.

Essex, M. J., Klein, M. H., Miech, R., & Smider, N. A. (2001). Timing of initial exposure to maternal major depression and children's mental health symptoms in kindergarten. *British Journal of Psychiatry, 179,* 151–156.

Farmer, A., Mahmood, A., Redman, K., Harris, T., Sadler, S., & McGuffin, P. (2003). A sib-pair study of the Temperament and Character Inventory scales in major depression. *Archives of General Psychiatry, 60,* 490–496.

Glover, V., & O'Connor, T. G. (2002). Editorial: Antenatal stress and later development: Findings, mechanisms, and further hypotheses. *British Journal of Psychiatry, 180,* 389–391.

Gunnar, M. R., Bruce, J., & Grotevant, H. D. (2000). In-

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

ternational adoption of institutionally reared children: Research and policy. *Development and Psychopathology, 12,* 677–697.

Haworth, G. (2001). Romanian adoptions in a worldwide context. *ACPP Occasional Papers, 18,* 45–50.

Hetherington, E. M. (Ed.). (1999). *Coping with divorce, single-parenting, and remarriage: A risk and resiliency perspective*. Mahwah, NJ: Erlbaum.

Hinshaw, S. P. (2002). Process, mechanism, and explanation related to externalizing behavior in developmental psychopathology. *Journal of Abnormal Child Psychology, 30,* 431–446.

Hodges, J., & Tizard, B. (1989). IQ and behavioural adjustment of ex-institutional adolescents. *Journal of Child Psychology and Psychiatry, 30,* 53–75.

Igumnov, S., & Drozdovitch, V. (2000). The intellectual development, mental and behavioural disorders in children from Belarus exposed in utero following the Chernobyl accident. *European Psychiatry, 15,* 244–253.

Kiernan, K. E. (1992). The impact of family disruption in childhood on transitions made in young adulthood. *Population Studies, 46,* 213–234.

Kiernan, K. E., & Hobcraft, J. (1997). Parental divorce during childhood: Age at first intercourse, partnership, and parenthood. *Population Studies, 51,* 41–55.

Kraemer, H. C., Stice, E., Kazdin, A., Offord, D., & Kupfer, D. (2001). How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. *American Journal of Psychiatry, 158,* 848–856.

Kreppner, J. M., O'Connor, T. G., Rutter, M., and the English and Romanian Adoptees Study Team. (2001). Can inattention/hyperactivity be a deprivation disorder? *Journal of Abnormal Child Psychology, 29,* 513–528.

Kurstjens, S., & Wolke, D. (2001). Effects of maternal depression on cognitive development of children over the first 7 years of life. *Journal of Child Psychology and Psychiatry, 42,* 623–636.

Lien, N. M., Meyer, K. K., & Winick, M. (1977). Early malnutrition and "late" adoption: A study of their effects on development of Korean orphans adopted into American families. *American Journal of Clinical Nutrition, 30,* 1734–1739.

McGlashan, T. H. (2001). Psychosis treatment prior to psychosis onset: Ethical issues. *Schizophrenia Research, 51,* 47–54.

Milan, S. E., & Pinderhughes, E. E. (2000). Factors influencing maltreated children's early adjustment in foster care. *Development and Psychopathology, 12,* 63–81.

Minty, B. (1999). Annotation: Outcomes in long-term foster family care. *Journal of Child Psychology and Psychiatry, 40,* 991–999.

Molitor, A., Mayes, L. C., & Ward, A. (2003). Emotion regulation behavior during a separation procedure in 18-month-old children of mothers using cocaine and other drugs. *Development and Psychopathology, 15,* 39–54.

Murray, L., Fiori–Cowley, A., & Hooper, R. (1996). The impact of postnatal depression and associated adversity on early mother–infant interactions and later infant outcome. *Child Development, 67,* 2512–2526.

Nyagu, A. I., Loganovsky, K. N., & Loganovskaja, T. K. (1998). Psychophysiologic aftereffects of prenatal irradiation. *International Journal of Psychophysiology, 30,* 303–311.

O'Connor, T. G., Dunn, J., Jenkins, J. M., Pickering, K., & Rasbash, J. (2001). Family settings and children's adjustment: Differential adjustment within and across families. *British Journal of Psychiatry, 179,* 110–115.

O'Connor, T. G., Marvin, R. S., Rutter, M., Olrick, J., Britner, P. A., & the E.R.A. Study Team. (2003). Child–parent attachment following early institutional deprivation. *Development and Psychopathology, 15,* 19–38.

O'Connor, T. G., Rutter, M., Beckett, C., Kreppner, J. M., & Keaveney, L. and the English and Romanian Adoptees Study Team. (2000). The effects of global severe privation on cognitive competence: Extension and longitudinal follow-up. *Child Development, 71,* 376–390.

O'Connor, T. G., Thorpe, K., Dunn, J., & Golding, J., and the ALSPAC Study Team (1999). Parental divorce and adjustment in adulthood: Findings from a community sample. *Journal of Child Psychology and Psychiatry, 40,* 777–790.

Reynolds, A. J. (1994). Effects of a preschool plus follow-on intervention for preschool children at risk. *Developmental Psychology, 30,* 787–804.

Rodgers, B., Power, C., & Hope, S. (1997). Parental divorce and adult psychological distress: Evidence from a national birth cohort: A research note. *Journal of Child Psychology and Psychiatry, 38,* 867–872.

Rushton, A., Treseder, J., & Quinton, D. (1995) An eight-year prospective study of older boys placed in permanent substitute families: A research note. *Journal of Child Psychology and Psychiatry 36,* 687–696.

Rutter, M., & the ERA Research Team (1998). Developmental catch-up and deficit following adoption after severe global early privation. *Journal of Child Psychology and Psychiatry, 39,* 465–476.

Rutter, M., O'Connor, T. G., & the ERA Research Team. (in press). Are there biological programming effects for psychological development? Findings from a study of Romanian adoptees. *Developmental Psychology*.

Rutter, M., Pickles, A., Murray, R., & Eaves, L. (2001). Testing hypotheses on specific environmental causal effects on behavior. *Psychological Bulletin, 127,* 291–324.

Sanders, M. R., Turner, K. M., & Markie-Dadds, C. (2002). The development and dissemination of the Triple P-Positive Parenting Program: A multilevel, evidence-based system of parenting and family support. *Prevention Science, 3,* 173–189.

Schneider, M. L., & Moore, C. F. (2000). Effect of prenatal stress on development: A nonhuman primate model. In C. Nelson (Ed.), *Minnesota Symposium on Child Psychology* (Vol. 30, pp. 201–243). Mahwah, NJ: Erlbaum.

Shealy, C. N. (1995). From Boys' Town to Oliver Twist: Separating fact from fiction in welfare reform and out-of-home placement of children and youth. *American Psychologist, 50,* 565–580.

Sinclair, D., & Murray, L. (1998). Effects of postnatal depression on children's adjustment to school. Teacher's reports. *British Journal of Psychiatry, 172,* 58–63.

Smotherman, W. P., & Robinson, S. R. (1990). The prenatal origins of behavioral organization. *Psychological Science, 1,* 97–106.

Thompson, W. R. (1957). Influence of prenatal maternal anxiety on emotionality in young rats. *Science, 125,* 698.

Tizard, B., & Rees, J. (1975). The effect of early institu-

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press

tional rearing on the behavioral problems and affectional relationships of four-year-old children. *Journal of Child Psychology and Psychiatry, 16,* 61–73.

Vaughn, B., Egeland, B., Sroufe, L. A., & Waters, E. (1979). Individual differences in infant–mother attachment at twelve and eighteen months: Stability and change in families under stress. *Child Development, 50,* 971–975.

Vondra, J. I., Shaw, D. S., Swearingen, L., Cohen, M., & Owens, E. B. (2001). Attachment stability and emotional and behavioral regulation from infancy to preschool age. *Development and Psychopathology, 13,* 13–33.

Vorria, P., Papaligoura, Z., Dunn, J., van IJzendoorn, M. H., Steele, H., Kontopoulou, A. & Sarafidou, Y. (in press). Early experiences and attachment relationships of Greek infants raised in residential group care. *Journal of Child Psychology and Psychiatry*.

Waters, E., Merrick, S., Treboux, D., Crowell, J., & Albersheim, L. (2000). Attachment security in infancy and early adulthood: A twenty-year longitudinal study. *Child Development, 71,* 684–689.

Watson, J. B., Mednick, S. A., Huttunen, M., & Wang, X. (1999). Prenatal teratogens and the development of adult mental illness. *Development and Psychopathology, 11,* 457–466.

Zeanah, C. H. (1996). Beyond insecurity: A reconceptualization of attachment disorders of infancy. *Journal of Consulting and Clinical Psychology, 64,* 42–52.

https://doi.org/10.1017/S0954579403000403 Published online by Cambridge University Press