Primate evidence on the late health effects of early-life adversity

Gabriella Conti[a,1], Christopher Hansman[b], James J. Heckman[a,c,d,e,1], Matthew F. X. Novak[f], Angela Ruggiero[f], and Stephen J. Suomi[f]

[a]The Harris School of Public Policy Studies and [c]Department of Economics, University of Chicago, Chicago, IL 60637; [b]Department of Economics, Columbia University, New York, NY 10027; [d]American Bar Foundation, Chicago, IL 60611; [e]University College Dublin, Dublin 4, Ireland; and [f]Section on Comparative Behavioral Genetics, National Institute of Child Health and Human Development, Poolesville, MD 20837

Contributed by James J. Heckman, April 10, 2012 (sent for review December 15, 2011)

This paper exploits a unique ongoing experiment to analyze the effects of early rearing conditions on physical and mental health in a sample of rhesus monkeys (*Macaca mulatta*). We analyze the health records of 231 monkeys that were randomly allocated at birth across three rearing conditions: mother rearing, peer rearing, and surrogate peer rearing. We show that the lack of a secure attachment relationship in the early years engendered by adverse rearing conditions has detrimental long-term effects on health that are not compensated for by a normal social environment later in life.

maternal behavior | social deprivation | long-term health

The importance of the early years in affecting a variety of aspects of later life, including health, through the biological embedding of early experiences is now widely recognized (1, 2). Some of the most compelling evidence on the consequences of early maternal and social deprivation comes from children raised in the adverse settings of Romanian orphanages of the 1980s and 1990s. Lasting physiological and mental effects have been striking there; the most recent findings (3) show a high degree of persistence until 15 y of age of quasiautism, disinhibited attachment, inattention/overactivity, and cognitive impairment. In addition, in the absence of malnutrition, children with institutional deprivation that lasted beyond the age of 6 mo had a major constraint in head growth. Orphans have been studied in a number of contexts. Ref. 4 is a meta-analysis of existing studies. Another environment in which children have been deprived of normal maternal relationships is the environment of the Israeli Kibbutz, where they were raised collectively. Ref. 5 provides a recent overview of the research on kibbutzim, concluding that children raised in this environment do not tend to strive for excellence, and the quality of their relationships when adults is diminished.

Apart from these atypical environments, the literature abounds with observational evidence on children who have been abused or neglected or have somehow not formed secure attachment relationships to their primary caregivers and have, subsequently, displayed maladaptive patterns of development (6, 7). Manipulating environments experimentally, however, is challenging with human populations (ref. 8 discusses recent exceptions). For decades, researchers have used nonhuman primate models to explore the behavioral and physiological effects of early maternal and social deprivation. Although the devastating social consequences of early isolation have been recognized since the work by Harlow and Zimmermann (9) in the 1950s and 1960s, more recent work has begun to uncover the impact of adverse rearing experiences on more direct physiological measures, including hormonal changes, brain function, and gene expression. For example, the work by Feng et al. (10) shows that the altered cortisol response to acute stressors in peer-reared monkeys is not reversed after 1.5 y of normal life. Additionally, the work by Spinelli et al. (11) finds that peer-reared monkeys display enlargement in stress-sensitive brain regions compared with mother-reared monkeys. The work by Jackowski et al. (12) documents that male bonnet monkeys subject to early stress show effects in brain development in the multiple regions involved in emotion processing, such as the corpus callosum and the hippocampus. However, despite the broad range of studies focusing on behavioral changes and physiological markers, relatively few studies have analyzed the health consequences of adverse early rearing conditions, focusing either on growth, reproduction, and survival (13) or cell-mediated immune response [e.g., Coe et al. (14) and Gordon et al. (15)]. In a recent summary, Schapiro (16) concludes that animate rather than inanimate enrichment (i.e., social housing rather than feeding enhancements) is effective in ameliorating the negative health consequences of adverse early conditions.

This paper contributes to the literature by exploiting experimental data on a sample of rhesus monkeys (*Macaca mulatta*)* subject to a randomized early rearing protocol to show evidence that the lack of a secure attachment relationship early in life has detrimental consequences on physical and mental health later in life. Furthermore, we show evidence that these effects differ by sex and stretch beyond the first year, suggesting that the consequences of early adversity get under the skin and are not compensated for by living in a normal social environment later in life.

PSYCHOLOGICAL AND COGNITIVE SCIENCES

Data

Our dataset is obtained from records collected until January of 2010 on 231 rhesus macaques born between 2002 and 2007 and raised in the Laboratory of Comparative Ethology, National Institute of Child Health and Human Development primate facilities at the National Institute of Health Animal Center as part of an ongoing randomized experiment. At birth, all subjects were randomized into one of three rearing conditions: mother reared (MR), peer reared (PR), and surrogate peer reared or surrogate reared (SPR). MR monkeys remained with their biological mothers from birth and were raised in large cages with other monkeys, whereas both PR and SPR were taken from their mothers and individually raised in a nursery until the 37th day of life. [Although all nursery reared (NR) monkeys are not breastfed, formula feeding cannot be considered the sole reason for our findings, because we observe differences between types of NR monkeys (PR vs. SPR).] On the 37th day, PR monkeys were placed in groups of four with the three monkeys closest in age. Monkeys in the same group spent 24 h together in a cage. They were removed only for testing. SPR monkeys spent 22 h/d alone in a cage with a surrogate mother (effectively a terry cloth-covered hot water bottle hanging from the top of the cage) and were placed with a peer group of three other SPR monkeys in a play cage that provided the opportunity for unlimited social interaction with the peers for the remaining 2 h each

Author contributions: G.C., J.J.H., and S.J.S. designed research; G.C., C.H., J.J.H., M.F.X.N., and A.R. performed research; G.C. and C.H. analyzed data; G.C., C.H., J.J.H., and S.J.S. wrote the paper; C.H., M.F.X.N., and A.R. assisted with data collection; and S.J.S. provided subjects.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. E-mail: gconti@uchicago.edu or jjh@uchicago.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1205340109/-/DCSupplemental.

*Although they are not our closest genetic relatives among nonhuman primates (they share about 95% of human genes, whereas chimpanzees and bonobos share 98–99%), they are like humans (and unlike virtually every other species of nonhuman primates) in their versatility and ability to adjust to and survive in almost any climate in the world. More information on the closeness between humans and macaques is in ref. 17.

Table 1. Primary outcomes: SPR vs. MR

Outcome	Effect			*p* values			
	Control mean	Unconditional	Conditional	Asymptotic	Naïve permutation	Conditional permutation	Conditional permutation (adj.)
Males							
Prevalence of stereotypy	0.215	0.716	0.669	**0.000**	**0.000**	**0.000**	**0.000**
Frequency of stereotypy	0.046	0.485	0.478	**0.000**	**0.000**	**0.000**	**0.000**
Prevalence of illness	0.723	0.208	0.177	**0.003**	**0.018**	**0.006**	**0.025**
Frequency of illness	0.154	0.135	0.120	**0.001**	**0.000**	**0.000**	**0.004**
Prevalence of wound	0.385	−0.005	−0.004	0.481	0.386	0.331	0.331
Frequency of wound	0.052	0.006	0.011	0.399	0.382	0.422	0.570
Prevalence of alopecia	0.369	0.217	0.156	**0.028**	**0.038**	**0.038**	0.150
Frequency of alopecia	0.113	0.053	0.041	0.106	**0.088**	**0.066**	0.387
Females							
Prevalence of stereotypy	0.070	0.756	0.656	**0.000**	**0.000**	**0.000**	**0.000**
Frequency of stereotypy	0.015	0.386	0.352	**0.000**	**0.000**	**0.000**	**0.000**
Prevalence of illness	0.649	0.220	0.157	**0.013**	**0.009**	**0.067**	0.210
Frequency of illness	0.161	0.082	0.075	**0.035**	**0.027**	**0.029**	0.160
Prevalence of wound	0.509	0.056	0.026	0.327	0.247	0.385	0.746
Frequency of wound	0.094	0.008	0.006	0.397	0.380	0.378	0.680
Prevalence of alopecia	0.403	0.249	0.116	**0.023**	**0.012**	0.181	0.322
Frequency of alopecia	0.137	0.035	0.005	0.202	0.187	0.522	0.522

$n = 94$ for males, and $n = 80$ for females. *p* values below 0.1 are in bold. Control is MR. Unconditional is unconditional difference in means between the treatment and control groups. The corresponding *p* values are computed in the columns asymptotic and naïve permutation. Conditional is conditional treatment effect with linear covariates year of birth and total time spent in the primate center. The corresponding *p* value is computed in the column conditional permutation. Asymptotic is one-sided *p* value for the hypothesis of no treatment effect based on asymptotic inference; the estimated effect size is in the unconditional column. Naïve permutation is the one-sided *p* value for the hypothesis of no treatment effect based on unconditional permutation inference; the estimated effect size is in the unconditional column. Conditional permutation is the one-sided *p* value for the hypothesis of no treatment effect based on the Freedman–Lane procedure using the linear covariates year of birth and total time spent in the primate center and restricting permutation orbits within strata formed by being a first- or later-born monkey; the estimated effect size is in the conditional column. Conditional permutation (adj.) is the *p* value from the previous column adjusted for multiple inferences using the step-down procedure.

day. Between 6 mo and 1 y, all monkeys born in the same year were put together in a single mixed social group (given that the average life of a monkey is 25 y in captivity, the period spent in treatment can be thought of as the critical 0–3 y in humans). MR monkeys constitute just over 50% of our sample, whereas PR and SPR monkeys make up just under 25% each (122, 57, and 52 monkeys, respectively); a slight majority of our sample is male (126 monkeys), and just under one-quarter of them were firstborn (56 monkeys) (Table S1 shows summary statistics).

The five outcomes analyzed in this paper are based on records from two sources: physical examinations and behavioral observations. Both were first performed at birth and continued throughout the lifecycle; both the examinations and observations are carried out uniformly across the rearing groups in our sample. Physical

Table 2. Primary outcomes: SPR vs. MR males

Outcome	Effect			*p* values			
	Control mean	Unconditional	Conditional	Asymptotic	Naïve permutation	Conditional permutation	Conditional permutation (adj.)
Main vs. other illness							
Prevalence of illness–main	0.723	0.001	−0.020	0.496	0.602	0.508	0.631
Frequency of illness–main	0.150	0.084	0.072	**0.028**	**0.017**	**0.031**	0.190
Prevalence of illness–other	0.200	0.352	0.311	**0.001**	**0.000**	**0.001**	**0.003**
Frequency of illness–other	0.029	0.072	0.066	**0.006**	**0.002**	**0.002**	**0.033**
Diarrhea vs. nondiarrhea illness							
Prevalence of illness–main	0.723	0.001	−0.020	0.496	0.602	0.508	0.631
Frequency of illness–main	0.150	0.084	0.072	**0.028**	**0.017**	**0.031**	0.190
Prevalence of illness–other: diarrhea	0.046	0.230	0.216	**0.007**	**0.003**	**0.004**	**0.033**
Frequency of illness–other: diarrhea	0.004	0.025	0.024	**0.015**	**0.001**	**0.003**	**0.053**
Prevalence of illness–other: nondiarrhea	0.154	0.225	0.196	**0.016**	**0.004**	**0.017**	0.129
Frequency of illness–other: nondiarrhea	0.025	0.047	0.042	**0.026**	**0.013**	**0.020**	0.179

$n = 94$ for males, and $n = 80$ for females. *p* values below 0.1 are in bold. Control is MR. Unconditional is unconditional difference in means between the treatment and control groups. The corresponding *p* values are computed in the columns asymptotic and naïve permutation. Conditional is conditional treatment effect with linear covariates year of birth and total time spent in the primate center. The corresponding *p* value is computed in the column conditional permutation. Asymptotic is one-sided *p* value for the hypothesis of no treatment effect based on asymptotic inference; the estimated effect size is in the unconditional column. Naïve permutation is the one-sided *p* value for the hypothesis of no treatment effect based on unconditional permutation inference; the estimated effect size is in the unconditional column. Conditional permutation is the one-sided *p* value for the hypothesis of no treatment effect based on the Freedman–Lane procedure using the linear covariates year of birth and total time spent in the primate center and restricting permutation orbits within strata formed by being a first- or later-born monkey; the estimated effect size is in the conditional column. Conditional permutation (adj.) is the *p* value from the previous column adjusted for multiple inferences using the step-down procedure. Notice that the multiple hypothesis testing correction also includes the other outcomes in Table 1. Full results are shown in Tables S2 and S3.

Table 3. Primary outcomes: PR vs. MR

Outcome	Effect			*p* value			
	Control mean	Unconditional	Conditional	Asymptotic	Naïve permutation	Conditional permutation	Conditional permutation (adj.)
Males							
Prevalence of stereotypy	0.215	0.441	0.470	**0.000**	**0.000**	**0.000**	**0.000**
Frequency of stereotypy	0.046	0.300	0.287	**0.000**	**0.000**	**0.000**	**0.000**
Prevalence of illness	0.723	−0.004	0.004	0.482	0.566	0.415	0.415
Frequency of illness	0.154	0.016	0.021	0.334	0.330	0.229	0.626
Prevalence of wound	0.385	0.053	0.042	0.313	0.389	0.380	0.692
Frequency of wound	0.052	0.045	0.044	0.111	**0.082**	**0.097**	0.478
Prevalence of alopecia	0.369	−0.057	0.018	0.291	0.221	0.367	0.602
Frequency of alopecia	0.113	−0.047	−0.030	**0.052**	**0.070**	0.253	0.502
Females							
Prevalence of stereotypy	0.070	0.770	0.751	**0.000**	**0.000**	**0.000**	**0.000**
Frequency of stereotypy	0.015	0.346	0.352	**0.000**	**0.000**	**0.000**	**0.000**
Prevalence of illness	0.649	0.071	0.058	0.264	0.192	0.318	0.318
Frequency of illness	0.161	−0.029	−0.027	0.202	0.218	0.257	0.503
Prevalence of wound	0.509	0.251	0.251	**0.013**	**0.008**	**0.017**	**0.046**
Frequency of wound	0.094	0.062	0.067	**0.054**	**0.028**	**0.028**	0.110
Prevalence of alopecia	0.403	0.316	0.261	**0.004**	**0.002**	**0.012**	**0.024**
Frequency of alopecia	0.137	0.116	0.104	**0.014**	**0.010**	**0.006**	**0.017**

$n = 97$ for males, and $n = 82$ for females. *p* values below 0.1 are in bold. Control is MR. Unconditional is unconditional difference in means between the treatment and control groups. The corresponding *p* values are computed in the columns asymptotic and naïve permutation. Conditional is conditional treatment effect with linear covariates year of birth and total time spent in the primate center. The corresponding *p* value is computed in the column conditional permutation. Asymptotic is one-sided *p* value for the hypothesis of no treatment effect based on asymptotic inference; the estimated effect size is in the unconditional column. Naïve permutation is the one-sided *p* value for the hypothesis of no treatment effect based on unconditional permutation inference; the estimated effect size is in the unconditional column. Conditional permutation is the one-sided *p* value for the hypothesis of no treatment effect based on the Freedman–Lane procedure using the linear covariates year of birth and total time spent in the primate center and restricting permutation orbits within strata formed by being a first- or later-born monkey; the estimated effect size is in the conditional column. Conditional permutation (adj.) is the *p* value from the previous column adjusted for multiple inferences using the step-down procedure.

examinations were performed four times per year by the facility veterinarians using a standardized worksheet. Items on this worksheet included physiological measures such as weight, a checklist for problems in various main bodily regions, and a space for descriptions of particular health issues not explicitly listed in a prespecified category. In our analysis, we examine three outcomes culled from these worksheets: a continuous measure of weight, binary indicators for health issues arising from wounds (wound), and health issues not caused by external bodily harm (illness). The category illness is constructed by including health issues recorded in two different parts of the primate physical health worksheet: problems in various main bodily regions [ears, eyes, nose, and throat (EENT), mouth/head, chest, abdomen, and urogenital] and issues recorded in the other section (the main categories here being diarrhea, rash, and hernia). The additional outcome measures that we analyze are binary indicators for the occurrence of any stereotypic behavior (stereotypy[†]) and the presence of hair loss (alopecia) obtained from 5-min focal points behavioral observations performed biannually by a skilled technician from the National Institute of Child Health and Human Development Research Animal Management Branch. For each dichotomous outcome reported, we construct two measures: one indicator of overall prevalence, which indicates whether the animal experienced the condition at any point during the period for which we have data available, and one indicator of frequency, which indicates the proportion of visits/observations in which the condition had been recorded. Illness and stereotypy had approximately the same frequency (18% and 21%, respectively), whereas alopecia is recorded, on average, in 14% of the observations, and wounds are recorded in 9% of the visits. In terms of overall prevalence, this amounts to 74% for illness, 48% for wound, and 46% for both stereotypy and alopecia (i.e., 172, 111, 106, and 107 monkeys have been recorded showing that particular condition at least one time during the full period of observation, respectively) (Table S1). When breaking down the category illness into its various components, we observe that 171 monkeys experience an illness because of problems in main bodily regions (henceforth, main illness) at least one time during the period for which we have data available, and the average frequency of main illness is 17%, and 71 monkeys experience an illness caused by other problems (henceforth, other illness), and the average frequency of other illness is 5%. Among the other illnesses, the numbers are fairly equally split between diarrhea and nondiarrhea-related conditions. Another breakdown shows that problems related to mouth and head are the most common pathologies in the main category and that rashes are most common after diarrhea in the other category.

Importantly, throughout our study, we only consider outcomes measured after the first year, when all monkeys have been placed into a common mixed social group, to study the long-term effects of adverse early rearing conditions. Additionally, we exploit supplementary data on intermediate phenotypes to try to understand the mechanisms—both physiological and behavioral—underlying the observed changes in later-life health outcomes to dig deeper and go beyond the estimates of average treatment effects.[‡]

Results

We present three main results on the later-life health effects of adverse early rearing conditions that are reported in Tables 1–3. We state and discuss each of them in turn.

[†]The full list of stereotypies observed includes digit sucking (the most frequent behavior), pacing, head tossing, self-grasping, saluting, spinning, rocking, circling, and swinging.

[‡]Although we base our analysis on a randomized experiment, we recognize the importance of understanding the mechanisms to accumulate useful knowledge, which can be used as a basis for implementation of policies (18). Inferences of causality based on increasing understanding over time of underlying mechanisms at the basis of observed effects are central to the process of knowledge accumulation. The successive developments and extensions of the Henle–Koch postulates and the corresponding changing guidelines for evaluating the causal role of an agent in infectious disease after technical developments in microbiology provide clear examples of the difficulties intrinsic to positing sufficient conditions for establishing causality (19).

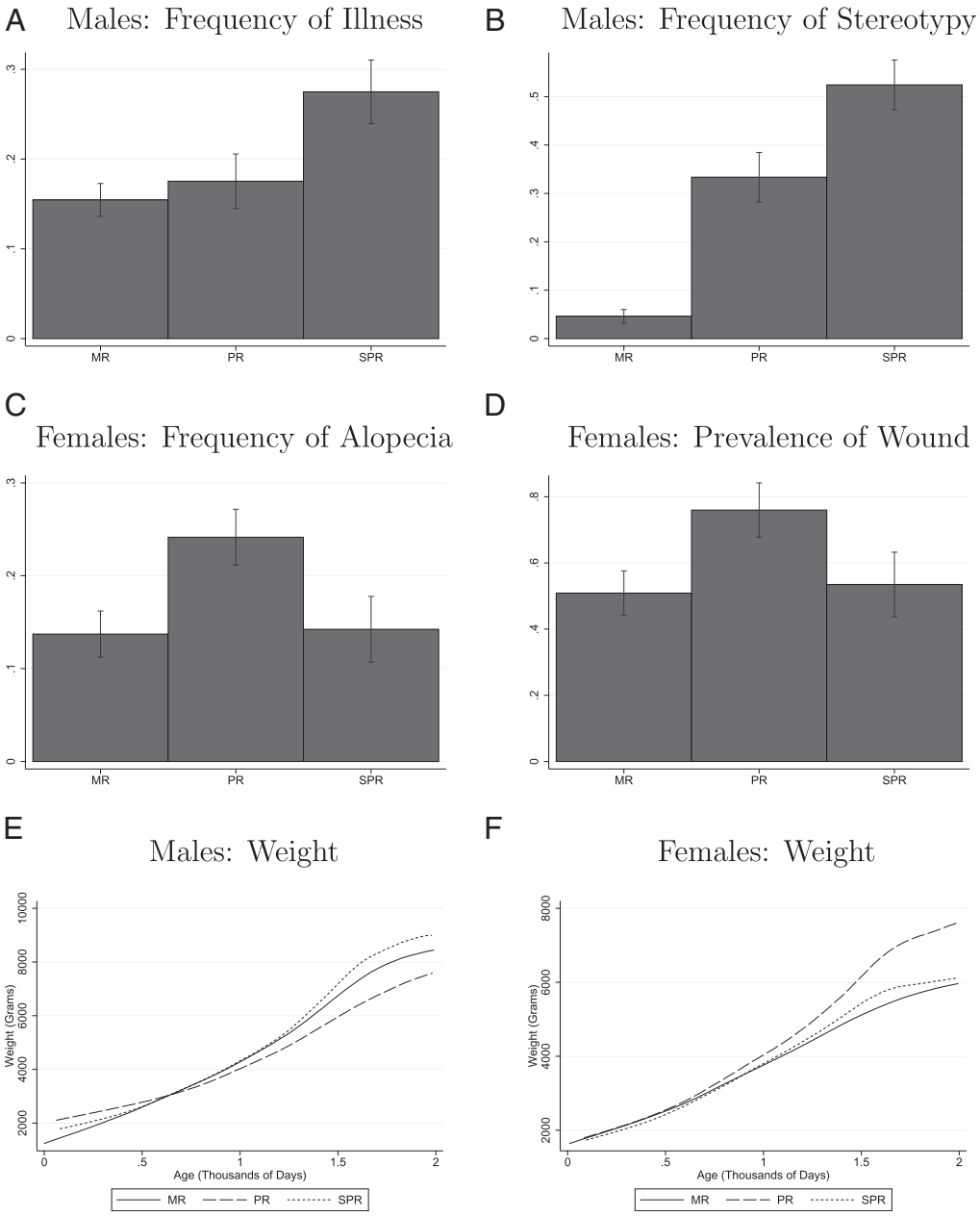

Fig. 1. Primary outcomes. (*A–D*) Predictions based on the results displayed in Tables 1 and 3. Standard error bars displayed. (*E* and *F*) Local polynomial regressions for weight over the lifecycle by rearing condition. Weight is measured in grams, and age is measured in thousands of days.

Physical Health: SPR Male Monkeys Exhibit a Statistically Significantly Higher Probability of Developing Illnesses in Terms of both Prevalence (p = 0.025) and Frequency (p = 0.004). Fig. 1*A* shows that the predicted frequency of illness for an SPR monkey is 0.274, which is almost two times as much as for an MR monkey (0.154; the value for SPR is obtained by summing the frequencies in the control mean and conditional columns in Table 1). The effect on prevalence is much more dramatic, with almost all of the SPR male monkeys having experienced an illness at least one time during the observation period. The adverse effects of SPR survive multiple hypothesis testing corrections as developed by ref. 20 and implemented in ref. 8, which can be seen in the conditional permutation (adj.) column in Table 1.

These results provide evidence of a causal link between early maternal and social deprivation and later-life illness. Our findings are consistent with the findings in ref. 21, which notes lifelong differences in cellular immune functioning and higher mortality rates among monkeys reared in isolation (with a protocol similar to the SPR). Additionally, we supplement our analysis with additional data on cortisol, adrenocortitropic hormone (ACTH), and 5-hydroxyindoleacetic acid (5-HIAA)[¶] collected while the monkeys are still in their respective treatment conditions (i.e., before 1 y of age). Our results are in line with the large body of observational evidence on humans on the role played by stress as a mediating factor between childhood adversity and later-life disease (22); we find higher cortisol levels among SPR male monkeys and deficits in both ACTH (23) and serotonin metabolism, because they have lower concentrations of 5-HIAA (the primary central serotonin metabolite), which is linked to aggression and antisocial behavior (24). Because they were collected while the monkeys were still in their separate rearing conditions, it is notable that visible health effects outlast this initial period. Notice that all of these effects survive corrections for multiple hypothesis testing (Table S4).

[¶]Because they were collected while the monkeys were still in their separate rearing conditions, it is notable that visible health effects outlast this initial period.

When breaking down the illness category into its various components (full results are presented in Tables S2 and S3) (Table 2), we notice that, although the adverse health effects of SPR seem pervasive (many outcomes are statistically significant when performing single hypothesis testing), only the effects on diarrhea survive multiple hypothesis testing corrections. Because of the importance of diarrhea [which could be caused by bacterial agents (*Campylobacter* and *Shigella* being the most common) or chronic in nature] in explaining the noted effects, we exploit additional information on medicines taken and blood test results for a small subsample of 34 monkeys to gain more insights into this condition. The medicines most commonly administered were erythromycin (used in the treatment of diarrhea caused by *Campylobacter*), metronidazole (used in the treatment of diarrhea caused by *Clostridium difficile*), baytril (used in the treatment of diarrhea caused by *Shigella*), and Imodium (used in the treatment of diarrhea caused by irritable bowel syndrome). Furthermore, the analysis of the blood test results (Table S3) reveals that male monkeys affected by diarrhea show abnormally lower values of sodium and potassium (as consequence of dehydration), abnormally higher values of blood urea nitrogen, and higher values of hematocrit and glucose (these values were much less altered for diarrhea-affected females).

Mental Health: NR (Both PR and SPR) Monkeys of both Genders Exhibit a Significantly Higher Probability of Developing Stereotypies in Terms of both Prevalence and Frequency ($p = 0.000$ for both Genders and Rearing Groups). Although the development of stereotypic behavior in response to adverse rearing conditions has been documented since the 1960s (25), recent research on humans gives a renewed importance to understanding this relationship. The work by Bos et al. (26) examines the connection between early institutionalization, foster care, and stereotypies in a cohort of Romanian children with a history of institutional care. The results establish evidence of an association between institutionalization and stereotypic behavior as well as an association between stereotypic behavior and both autism and cognitive and language deficits. They establish the potential for remediation through foster care. In light of their findings, our results serve to not only highlight the parallels between NR in monkeys and institutionalization in human infants but also show the power of adverse early experience to produce behavioral abnormalities that are, at the very least, markers of deeper developmental deficits. It should also be noted that, in males, SPR produces a significantly higher frequency of stereotypies compared with PR (Fig. 1*B*), whereas in females, the two rearing statuses are not significantly differently affected.

PR Female Monkeys Exhibit a Significantly Higher Probability of Being Wounded ($p = 0.046$) and Experiencing Alopecia in Terms of both Prevalence ($p = 0.024$) and Frequency ($p = 0.017$), and They Have a Significantly Higher Weight than Their MR or SPR Counterparts ($p = 0.043$). Hence, as shown in Fig. 1 *C* and *D*, it seems that female monkeys raised with males in mixed sex groups develop patterns of behavior that are convergent with those patterns of males. Again, we supplement our analysis with additional data collected while the monkeys were still in their separate rearing conditions to investigate the early behavioral origins of these later-life differences. Given the external nature of wounds and alopecia, we examine whether behavioral differences across monkeys allocated to different rearing conditions as opposed to physiological changes could account for these effects. We find that PR females display higher levels of aggressive behavior compared with MR females (evidence presented in Table S5), suggesting that alopecia might be partly because of hair pulling by others [the work by Novak and Meyer (27) shows that it can be caused by a variety of factors, including nutritional imbalances and hair pulling by others], and contrary to SPR females, they did not show self-grooming (which includes self-scratching or biting) behavior, suggesting that the wounds recorded during the physical examination are likely not caused by self-harm.[§] A similar reduction in the typical sexual dimorphism is observed with respect to weight: the weight of PR animals of both sexes effectively converges (Fig. 1 *E* and *F* and Table S6). The unconditional mean difference between the weight of male and female PR monkeys in our sample is 25 g, which is not statistically significantly different from zero; we do observe, instead, statistically significant sex differences between MR (495 g) and SPR (1,099 g) monkeys. We interpret this convergence pattern as the result of the influence of the early social rearing environment (peer groups are of mixed sex) on the expression of behavioral sex differences (ref. 29 is a review of the role of nature and nurture on the development of sexually dimorphic behavior in rhesus monkeys). We observe converging patterns for both male and female monkeys, and we do not find statistically significant evidence of a stress-related response (no statistically significant differences in cortisol levels, ACTH, and 5-HIAA concentration in the female PR monkeys compared with the MR monkeys, which is shown in Table S4).[‖] Although sex differences in the effects of early-life experiences on behavior have already been reported in the literature [earlier studies (31) show that males are much more affected by early deprivation than females], our study documents differential response by sex with respect to health outcomes.

Methods

The strength of our research design stems from the experimental manipulation of the early environment, because the monkeys are randomly allocated at birth across different rearing conditions. In this way, we exploit the major benefit of randomization, which avoids the problem of selection bias [i.e., ensuring that $(Y_0, Y_1) \perp\!\!\!\perp D$, where D is the treatment assignment indicator (where monkeys can be assigned to either the PR or SPR condition), $\perp\!\!\!\perp$ stands for statistical independence, and (Y_0, Y_1) are vectors of potential outcomes for treated and control units].** However, the benefits of randomization in terms of protection against bias from unknown potentially influential factors are lost when the allocation of participants to treatment and control units is compromised (i.e., when treatments and controls have imbalanced covariate distribution). In our case, this result occurs for two reasons: there is an imbalance of rearing statuses across cohorts (older cohorts are more likely to be SPR), and first-born monkeys are preferentially kept with mothers according to laboratory protocol. The assumption of independence between potential outcomes and treatment assignment has to be modified to read $(Y_0, Y_1) \perp\!\!\!\perp D|X$, where X is the year of birth and primiparousness. Because we have knowledge of the variables that determine assignment to treatment, we can match on them to account for departures from the randomization protocol.

Our aim is to test the null hypothesis of no effect of PR and SPR treatment conditions on several later-life outcomes, which can be formally stated as (1)

$$Y \perp\!\!\!\perp D|X \qquad [1]$$

where Y is the outcome vector, $D = 0$ if MR, and $D = 1$ if PR or SPR, respectively. However, our small sample size calls into question the validity of applying classical tests based on large-sample statistical theory. Hence, we use permutation-based inference as an alternative approach; we perform one-sided permutation tests (we allow for unequal variances across the

[§]The work by Lutz et al. (28) also reports that self-biting is more common among SPR than MR or PR monkeys.

[‖]Another possible explanation for these findings would be related to the dynamics of social hierarchy after the transition from the respective rearing environments to the common social group and its relation to weight gain (30). Unfortunately, the current unavailability of social dominance data and information on food consumption and stress-related biomarkers after the end of the treatment prevents us from assessing the plausibility of this explanation; therefore, we defer the answer to this question to another occasion.

**Ref. 32 has a thorough discussion of the scientific model of causality. The standard model of program evaluation describes the observed outcome for participant i, Y_i, by $Y_i = D_i Y_{i,1} + (1 - D_i) Y_{i,0}$, where $(Y_{i,0}, Y_{i,1})$ are potential outcomes corresponding to control and treatment status for participant i, respectively, and D_i is the assignment indicator ($D_i = 1$ if treated and $D_i = 0$ otherwise). An evaluation problem occurs, because either $Y_{i,0}$ or $Y_{i,1}$ is observed but not both. Properly designed and implemented randomized experiments can eliminate this problem for mean treatment effects, because they produce independence between $(Y_{i,0}, Y_{i,1})$ and D_i. Within this setup, we refer throughout our analysis to treatment as the PR and SPR conditions and to control as the MR condition.

groups) applying the Freedman–Lane procedure, using as covariates year of birth and total time spent in the primate center, and restricting the permutation orbits within strata formed by being first or later born.[††] Additionally, we consider several prevalence and frequency measures. To avoid the problems of multiple hypothesis testing and selecting singly significant results from a set of largely statistically insignificant outcomes, we control for multiple hypothesis testing using the step-down procedure developed in the work by Romano and Wolf (20), as implemented in ref. 8.[‡‡]

[††]This approach is the approach that was used in the evaluation of the Perry Preschool Program in ref. 8. It involved testing a null hypothesis (i.e., the hypothesis that the experiment had no impact) using permutations of the data. Taking permutations of the data means randomly switching the treatment assignment of the monkeys between the MR and the PR and SPR conditions, respectively. The null hypothesis of no treatment effect is equivalent to the statement that the distribution of the outcomes of the treatment and control groups is the same. We used 10,000 permutations throughout.

[‡‡]This method corrects for multiple hypothesis testing using the family-wise error rate (i.e., the probability of obtaining one or more false positives out of a set of hypotheses tests). The work by Romano and Wolf (20) has shown that this step-down procedure exhibits strong family-wise error rate control, and it is less conservative than traditional procedures and obtains gain in power from accounting for statistical dependencies among the test statistics associated with each individual hypothesis. *Methods* has details on the implementation of the procedures that we adopt.

Conclusions

Although the importance of the early years of life in affecting adult outcomes is now recognized, establishing the existence of a causal effect on health of early exposure to adversity can be a challenging task. In this paper, we exploit a unique ongoing experiment in a colony of rhesus monkeys to provide causal evidence of the health effects of early maternal and social deprivation. We show that the lack of a secure attachment relationship in the early years has detrimental consequences for both physical and mental health later in life, with long-lasting effects that vary by sex. The persistence of these effects after the end of the treatment emphasizes the need to intervene early in life to prevent long-term damage.

ACKNOWLEDGMENTS. We thank Ruth A. Woodward, DVM, for providing the medical records and Ernest B. Davis for providing behavioral observations. We also thank Meiping Sun for her help in the data collection in the first stage of this process. This work was supported in part by the American Bar Foundation, the J. B. and M. K. Pritzker Family Foundation, the Susan Thompson Buffett Foundation, National Institute of Child Health and Human Development Grant R37-HD065072, National Institute on Aging Grant R01-AG034679, a project with the Becker Friedman Institute for Research and Economics funded by the Institute for New Economic Thinking, and an anonymous funder.

1. Fox SE, Levitt P, Nelson CA III (2010) How the timing and quality of early experiences influence the development of brain architecture. *Child Dev* 81:28–40.
2. Hertzman C, Boyce T (2010) How experience gets under the skin to create gradients in developmental health. *Annu Rev Public Health* 31:329–347.
3. Rutter M (2010) Deprivation-specific psychological patterns: Effects of institutional deprivation. *Monogr Soc Res Child Dev* 75:1–242.
4. Bakermans-Kranenburg MJ, van Ijzendoorn MH, Juffer F (2008) Earlier is better: A meta-analysis of 70 years of intervention improving cognitive development in institutionalized children. *Monogr Soc Res Child Dev* 73:279–293.
5. Lieblich A (2010) A century of childhood, parenting, and family life in the kibbutz. *J Isr Hist* 29:1–24.
6. Cicchetti D, Toth SL (1995) A developmental psychopathology perspective on child abuse and neglect. *J Am Acad Child Adolesc Psychiatry* 34:541–565.
7. Nunes S, Watanabe M, Morimoto H, Moriya R, Reiche E (2010) The impact of childhood sexual abuse on activation of immunological and neuroendocrine response. *Aggress Violent Behav* 15:440–445.
8. Heckman JJ, Moon SH, Pinto R, Savelyev PA, Yavitz AQ (2010) Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quant Econom* 1:1–46.
9. Harlow H, Zimmermann R (1959) Affectional responses in the infant monkey. *Science* 130:419–467.
10. Feng X, et al. (2011) Maternal separation produces lasting changes in cortisol and behavior in rhesus monkeys. *Proc Natl Acad Sci USA* 108:14312–14317.
11. Spinelli S, et al. (2009) Early-life stress induces long-term morphologic changes in primate brain. *Arch Gen Psychiatry* 66:658–665.
12. Jackowski A, et al. (2011) Early-life stress, corpus callosum development, hippocampal volumetrics, and anxious behavior in male nonhuman primates. *Psychiatry Res* 192:37–44.
13. Sackett GP, Ruppenthal GC, Davis AE (2002) Survival, growth, health, and reproduction following nursery rearing compared with mother rearing in pigtailed monkeys (*Macaca nemestrina*). *Am J Primatol* 56:165–183.
14. Coe CL, Lubach GR, Ershler WB, Klopp RG (1989) Influence of early rearing on lymphocyte proliferation responses in juvenile rhesus monkeys. *Brain Behav Immun* 3: 47–60.
15. Gordon TP, et al. (1992) Social separation and reunion affects immune system in juvenile rhesus monkeys. *Physiol Behav* 51:467–472.
16. Schapiro SJ (2002) Effects of social manipulations and environmental enrichment on behavior and cell-mediated immune responses in rhesus macaques. *Pharmacol Biochem Behav* 73:271–278.
17. Maestripieri D (2007) *Macachiavellian Intelligence: How Rhesus Macaques and Humans Have Conquered the World* (University of Chicago Press, Chicago).
18. Heckman JJ, Malofeeva L, Pinto R, Savelyev PA (2012) Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *Am Econ Rev*, in press.
19. Evans AS (1976) Causation and disease: The Henle-Koch postulates revisited. *Yale J Biol Med* 49:175–195.
20. Romano JP, Wolf M (2005) Exact and approximate stepdown methods for multiple hypothesis testing. *J Am Stat Assoc* 100:94–108.
21. Lewis MH, Gluck JP, Petitto JM, Hensley LL, Ozer H (2000) Early social deprivation in nonhuman primates: Long-term effects on survival and cell-mediated immunity. *Biol Psychiatry* 47:119–126.
22. McEwen BS (1998) Protective and damaging effects of stress mediators. *N Engl J Med* 338:171–179.
23. Vasterling JJ, Brewin CR, eds (2005) *Neuropsychology of PTSD: Biological, Cognitive, and Clinical Perspectives* (The Guilford Press, New York).
24. Higley JD, et al. (1996) CSF testosterone and 5-HIAA correlate with different types of aggressive behaviors. *Biol Psychiatry* 40:1067–1082.
25. Harlow HF, Harlow MK (1962) The effect of rearing conditions on behavior. *Bull Menninger Clin* 26:213–224.
26. Bos KJ, Zeanah CH, Jr., Smyke AT, Fox NA, Nelson CA III (2010) Stereotypies in children with a history of early institutional care. *Arch Pediatr Adolesc Med* 164:406–411.
27. Novak MA, Meyer JS (2009) Alopecia: Possible causes and treatments, particularly in captive nonhuman primates. *Comp Med* 59:18–26.
28. Lutz CK, Davis EB, Ruggiero AM, Suomi SJ (2007) Early predictors of self-biting in socially-housed rhesus macaques (*Macaca mulatta*). *Am J Primatol* 69:584–590.
29. Wallen K (1996) Nature needs nurture: The interaction of hormonal and social influences on the development of behavioral sex differences in rhesus monkeys. *Horm Behav* 30:364–378.
30. Bastian ML, Sponberg AC, Sponberg AC, Suomi SJ, Higley JD (2003) Long-term effects of infant rearing condition on the acquisition of dominance rank in juvenile and adult rhesus macaques (*Macaca mulatta*). *Dev Psychobiol* 42:44–51.
31. Sackett GP (1972) Exploratory behavior of rhesus monkeys as a function of rearing experiences and sex. *Dev Psychol* 6:260–270.
32. Heckman JJ (2008) Econometric causality. International Statistical Review, 76(1):1–27.

Supporting Information

Conti et al. 10.1073/pnas.1205340109

SI Methods

1.1 Permutation. The permutation testing procedure relies on the exchangeability properties of the joint distribution of outcomes and treatment assignments (i.e., the joint distribution is invariant to permutation of its elements). In practice, the permutation testing procedure compares a test statistic computed on the unpermuted data with a distribution of test statistics computed by resampling the data. In cases like our study, however, where randomization has been compromised, conditional inference can be implemented. In this case, the conditional exchangeability property is applied, and independence between the distribution of outcomes and treatment assignment is tested conditional on a set of variables (X). Given our sample size, full nonparametric conditioning is not an option, and therefore, we assumed a linear relationship between the outcomes and a subset of the variables (X^L; year of birth and total time spent in the primate center), and we restricted the permutation orbits to the remaining subset (X^N; a dichotomous indicator for being a firstborn monkey). According to this procedure, residuals computed from a regression of the outcomes on the covariates for which a linear relationship is assumed (X^L) are permuted within orbits defined by the variables that enter nonparametrically (X^N). This method is known as the Freedman–Lane (1) procedure, and it has been found to be superior to the others in a series of Monte Carlo studies (2).

1.2 Step-Down Procedure. To illustrate the step-down procedure, consider the null hypothesis of no treatment effect for a set of K joint outcomes, where the complement of this set is that there exists at least one hypothesis of K that we reject. The procedure by Romano and Wolf (3), as adapted to the current setting by ref. 4 starts by considering a joint test of all null hypotheses for the set of K hypotheses by comparing the maximum of the set of statistics associated with the hypotheses being jointly tested with the α-quantile of its distribution (α is the level of family-wise error rate for which we want to control) to determine whether this first joint hypothesis is rejected or not. If we fail to reject the joint null hypothesis, then the algorithm stops; if we reject it, then we iterate and consider successive joint hypotheses that exclude the outcomes with the highest associated test statistics. Therefore, the procedure steps down, and at each successive step, it is implemented on a set of $K - 1$ null hypotheses. The process iterates until only one hypothesis remains.

1. Freedman D, Lane D (1983) A nonstochastic interpretation of reported significance levels. *J Bus Econ Stat* 1:292–298.
2. Anderson MJ, Legendre P (1999) An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. *J Stat Comput Simul* 62: 271–303.
3. Romano JP, Wolf M (2005) Exact and approximate stepdown methods for multiple hypothesis testing. *J Am Stat Assoc* 100:94–108.
4. Heckman JJ, Moon SH, Pinto R, Savelyev PA, Yavitz AQ (2010) Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. Quant Econom 1:1–46.

Table S1. Summary statistics

	Subjects	Percentage (%)
Rearing status		
Mother reared	122	52.81
Peer reared	57	24.68
Surrogate peer reared	52	22.51
Sex		
Male	126	54.55
Female	105	45.55
Birth order		
Primiparious	56	24.24
Multiparious	175	75.76
Year of birth		
2002	41	17.75
2003	31	13.42
2004	31	13.42
2005	38	16.45
2006	46	19.91
2007	44	19.05
Outcomes	Prevalence	Frequency
Stereotypy	0.46	0.21
Illness	0.74	0.18
Illness–main	0.74	0.17
Illness–main: EENT	0.22	0.02
Illness–main: mouth/head	0.50	0.08
Illness–main: abdominal	0.32	0.05
Illness–main: chest	0.13	0.02
Illness–main: urogenital	0.16	0.03
Illness–other	0.31	0.05
Illness–other: diarrhea	0.14	0.02
Illness–other: nondiarrhea	0.19	0.03
Illness–other nondiarrhea: rash	0.10	0.02
Wound	0.48	0.09
Alopecia	0.46	0.14
	Observations	Mean (SD)
Weight (g)	2,636	4,453 (2,113)

All summary statistics refer to the analytical sample of 231 monkeys observed after their first year of life. The category illness–main includes ears, eyes, nose, and throat (EENT), mouth/head, abdominal, chest, and urogenital issues. The category illness–other includes both diarrhea- and nondiarrhea-related diseases (rash being the bigger component of the latter).

Table S2. Primary outcomes: Surrogate peer reared vs. mother reared, males

Outcome	Effect			*p* values			
	Control mean	Unconditional	Conditional	Asymptotic	Naïve permutation	Conditional permutation	Conditional permutation (adj.)
Main vs. other illness							
Prevalence of stereotypy	0.215	0.716	0.669	**0.000**	**0.000**	**0.000**	**0.000**
Frequency of stereotypy	0.046	0.485	0.478	**0.000**	**0.000**	**0.000**	**0.000**
Prevalence of illness–main	0.723	0.001	−0.020	0.496	0.602	0.508	0.631
Frequency of illness–main	0.150	0.084	0.072	**0.028**	**0.017**	**0.031**	0.190
Prevalence of illness–other	0.200	0.352	0.311	**0.001**	**0.000**	**0.001**	**0.003**
Frequency of illness–other	0.029	0.072	0.066	**0.006**	**0.002**	**0.002**	**0.033**
Prevalence of wound	0.385	−0.005	−0.004	0.481	0.386	0.331	0.331
Frequency of wound	0.052	0.006	0.011	0.399	0.382	0.422	0.719
Prevalence of alopecia	0.369	0.217	0.156	**0.028**	**0.038**	**0.038**	0.210
Frequency of alopecia	0.113	0.053	0.041	0.106	**0.088**	**0.066**	0.514
Diarrhea vs. nondiarrhea illness							
Prevalence of stereotypy	0.215	0.716	0.669	**0.000**	**0.000**	**0.000**	**0.000**
Frequency of stereotypy	0.046	0.485	0.478	**0.000**	**0.000**	**0.000**	**0.000**
Prevalence of illness–main	0.723	0.001	−0.020	0.496	0.602	0.508	0.631
Frequency of illness–main	0.150	0.084	0.072	**0.028**	**0.017**	**0.031**	0.190
Prevalence of illness–other: diarrhea	0.046	0.230	0.216	**0.007**	**0.003**	**0.004**	**0.033**
Frequency of illness–other: diarrhea	0.004	0.025	0.024	**0.015**	**0.001**	**0.003**	**0.053**
Prevalence of illness–other: nondiarrhea	0.154	0.225	0.196	**0.016**	**0.004**	**0.017**	0.129
Frequency of illness–other: nondiarrhea	0.025	0.047	0.042	**0.026**	**0.013**	**0.020**	0.179
Prevalence of wound	0.385	−0.005	−0.004	0.481	0.386	0.331	0.331
Frequency of wound	0.052	0.006	0.011	0.399	0.382	0.422	0.719
Prevalence of alopecia	0.369	0.217	0.156	**0.028**	**0.038**	**0.038**	0.210
Frequency of alopecia	0.113	0.053	0.041	0.106	**0.088**	**0.066**	0.514

$n = 94$. *p* values below 0.1 are in bold. Control is mother reared. Unconditional is the unconditional difference in means between the treatment and control groups. The corresponding *p* values are computed in the columns asymptotic and naïve permutation. Conditional is the conditional treatment effect with linear covariates year of birth and total time spent in the primate center. The corresponding *p* value is computed in the column conditional permutation. Asymptotic is the one-sided *p* values for the hypothesis of no treatment effect based on asymptotic inference; the estimated effect size is in the unconditional column. Naïve permutation is the one-sided *p* value for the hypothesis of no treatment effect based on unconditional permutation inference; the estimated effect size is in the unconditional column. Conditional permutation is the one-sided *p* value for the hypothesis of no treatment effect based on the Freedman–Lane (1) procedure using the linear covariates year of birth and total time spent in the primate center and restricting permutation orbits within strata formed by being a first- or later-born monkey; the estimated effect size is in the conditional column. Conditional permutation (adj.) is the *p* value from the previous column adjusted for multiple inferences using the step-down procedure.

1. Freedman D, Lane D (1983) A nonstochastic interpretation of reported significance levels. *J Bus Econ Stat* 1:292–298.

Table S3. Auxiliary outcomes and blood tests

	Effect		*p* values	
Outcome	Control mean	Treatment effect	Permutation	Permutation (adj.)
Males				
Potassium	3.900	−0.988	**0.000**	**0.001**
Sodium	144.583	−8.083	**0.002**	**0.003**
Hematocrit	37.858	5.873	**0.001**	**0.002**
Glucose	79.333	33.260	**0.007**	**0.007**
Blood urea nitrogen	16.250	16.906	**0.000**	**0.001**
Potassium (abnormal)	0.000	0.438	**0.000**	**0.002**
Sodium (abnormal)	0.083	0.385	**0.002**	**0.003**
Hematocrit (abnormal)	0.583	0.104	0.377	0.535
Glucose (abnormal)	0.083	0.010	0.265	0.265
Blood urea nitrogen (abnormal)	0.000	0.313	**0.000**	**0.003**
Females				
Potassium	3.613	−0.726	**0.000**	**0.019**
Sodium	141.813	−5.419	**0.013**	**0.026**
Hematocrit	29.531	13.204	**0.000**	**0.000**
Glucose	91.375	5.155	0.365	0.591
Blood urea nitrogen	36.625	−2.973	0.334	0.334
Potassium (abnormal)	0.000	0.424	**0.000**	**0.000**
Sodium (abnormal)	0.313	0.263	**0.017**	0.104
Hematocrit (abnormal)	0.500	0.333	**0.008**	**0.082**
Glucose (abnormal)	0.188	0.070	0.184	0.424
Blood urea nitrogen (abnormal)	0.375	0.034	0.302	0.302

For each treatment–control comparison, we present two sets of results: one set using the absolute value and another set using binary indicators for the presence of abnormal values. Normal ranges for the various blood tests are as follows: sodium = 140–160 mEq/L; potassium = 2.3–6.7 mEq/L; blood urea nitrogen = 8.0–30.0 mg/dL; glucose = 60–160 g/dL; and hematocrit = 30–38%. $n = 44$ for males; $n = 82$ for females. *p* values below 0.1 are in bold. Treatment effect is the unconditional difference in means between the treatment and control groups. Notice here that the treatment group is the one affected by diarrhea. The corresponding *p* values are computed in the column permutation. Permutation is the one-sided *p* value for the hypothesis of no treatment effect based on unconditional permutation inference; the estimated effect size is in the treatment effect column. Permutation (adj.) is the *p* value from the previous column adjusted for multiple inferences using the step-down procedure.

Table S4. Auxiliary outcomes

Outcome	Effect			*p* values		
	Control mean	Unconditional	Conditional	Naïve permutation	Conditional permutation	Conditional permutation (adj.)
Surrogate peer reared vs. mother reared males						
Cortisol (μg/dL) mean	50.926	7.889	10.296	**0.047**	**0.019**	**0.022**
ACTH (pg/mL) mean	221.677	−56.356	−43.647	**0.023**	**0.023**	**0.023**
5-HIAA (pmol/mL) mean	716.824	−112.011	−93.850	**0.002**	**0.065**	**0.069**
Cortisol (μg/dL) day 90	50.671	10.797	11.932	**0.014**	**0.016**	**0.027**
ACTH (pg/mL) day 90	212.143	−45.191	−43.512	**0.092**	**0.040**	**0.040**
5-HIAA (pmol/mL) day 90	675.786	−79.632	−71.269	**0.018**	**0.061**	**0.067**
Surrogate peer reared vs. mother reared females						
Cortisol (μg/dL) mean	56.440	−1.595	−1.847	0.308	0.275	0.466
ACTH (pg/mL) mean	269.795	−72.477	−70.491	0.107	**0.092**	0.234
5-HIAA (pmol/mL) mean	706.200	−54.109	−29.648	0.303	0.346	0.346
Cortisol (μg/dL) day 90	55.622	−1.667	−1.801	0.352	0.344	0.625
ACTH (pg/mL) day 90	263.489	−85.289	−88.818	**0.075**	**0.070**	0.191
5-HIAA (pmol/mL) day 90	608.667	0.444	15.489	0.492	0.354	0.354
Peer reared vs. mother reared males						
Cortisol (μg/dL) mean	50.926	2.004	3.960	0.316	0.211	0.326
ACTH (pg/mL) mean	221.677	11.134	−12.352	0.406	0.362	0.362
5-HIAA (pmol/mL) mean	716.824	−135.824	−71.565	**0.035**	0.147	0.270
Cortisol (μg/dL) day 90	50.671	3.329	4.654	0.285	0.279	0.388
ACTH (pg/mL) day 90	212.143	−2.493	−30.038	0.500	0.253	0.253
5-HIAA (pmol/mL) day 90	675.786	−138.036	−59.694	0.046	0.122	0.296
Peer reared vs. mother reared females						
Cortisol (μg/dL) mean	56.440	7.197	6.639	0.247	0.259	0.432
ACTH (pg/mL) mean	269.795	−44.032	−59.587	0.282	0.509	0.575
5-HIAA (pmol/mL) mean	706.200	−70.450	−21.428	0.367	0.395	0.395
Cortisol (μg/dL) day 90	55.622	16.078	13.371	0.298	0.404	0.737
ACTH (pg/mL) day 90	263.489	−29.339	−23.894	0.416	0.805	0.805
5-HIAA (pmol/mL) day 90	608.667	−123.167	−42.167	**0.070**	0.341	0.804

Blood (assayed for ACTH and cortisol) and cerebrospinal fluid (assayed for concentrations of the 5-HT metabolite 5-HIAA with gas chromatography–MS) were collected in days 60, 90, 120, and 150 in the first 5 mo of life of the monkeys up until 2005; therefore, they are only available for a subsample. For each treatment–control comparison, we present two sets of results: one set using the mean value (in case there is more than one valid observation for each measurement) and another set using the day 90 measurements (the day in which the sample size is maximized). We exclude the day 60 cortisol and adrenocortitropic hormone (ACTH) measurements, because the month 2 sample was taken during a nonstressed session, whereas the other samples were collected after a 30-min separation and isolation period in a single 64 × 61 × 76-cm cage in an empty room. $n = 33$ (mean) and $n = 27$ (day 90) in the surrogate peer reared vs. mother reared males group, $n = 21$ (mean) and $n = 18$ (day 90) in the surrogate peer reared vs. mother reared females group, $n = 22$ (mean) and $n = 18$ (day 90) in the peer reared vs. mother reared males group, and $n = 14$ (mean) and $n = 11$ (day 90) in the peer reared vs. mother reared females group. *p* values below 0.1 are in bold. Unconditional is the unconditional difference in means between the treatment and control groups. The corresponding *p* values are computed in the naïve permutation column. Conditional is the conditional treatment effect with linear covariate year of birth. The corresponding *p* value is computed in the column conditional permutation. Naïve permutation is the one-sided *p* value for the hypothesis of no treatment effect based on unconditional permutation inference; the estimated effect size is in the unconditional column. Conditional permutation is the one-sided *p* value for the hypothesis of no treatment effect based on the Freedman–Lane (1) procedure using the linear covariate year of birth and restricting permutation orbits within strata formed by being the first- or later-born monkey; the estimated effect size is in the conditional column. Conditional permutation (adj.) is the *p* value from the previous column adjusted for multiple inferences using the step-down procedure. 5-HIAA, 5-hydroxyindoleacetic acid.

1. Freedman D, Lane D (1983) A nonstochastic interpretation of reported significance levels. *J Bus Econ Stat* 1:292–298.

Table S5. Aggression and self-grooming in females

	Aggression	Self-grooming
Peer reared	0.038* (0.013)	0.074 (0.049)
Surrogate peer reared	−0.002 (0.005)	0.366* (0.062)
Observations	1,912	1,911

The numbers above are coefficients from linear regressions. Aggression and self-grooming are binary variables for the existence of such behaviors in a 5-min observation period. They were recorded two times per week in the first 30 wk of life of the monkeys when they were still in their separate rearing conditions. The numbers in parentheses are robust SEs clustered at the individual level. Peer reared and surrogate peer reared are binary indicators of the respective rearing statuses. Controls for year of birth, primiparous birth and of measurement week are also included in each specification.
*$p < 0.01$.

Table S6. Primary outcome: Weight

	Male	Female
PR	−99.051 (98.838)	474.985* (231.737)
SPR	92.234 (173.340)	−152.025 (174.723)
Observations	1,420	1,216

Weight is a continuous variable measured in grams. Included above are the coefficients δ_1 and δ_2 from a linear regression estimated by ordinary least squares of the following form: $Y_{i,t} = \alpha + \delta_1 D_{i,PR} + \delta_2 D_{i,SPR} + \beta X_{i,t} + e_{i,t}$, where $Y_{i,t}$ is the weight of monkey i at time t, $D_{i,PR}$ and $D_{i,SPR}$ are two dummies for treatment status (we set MR as the baseline), and X is a set of basic controls that includes dummies for year of birth, a binary indicator for first-born status, and age at the time of the examination. Included in parentheses are robust SEs clustered at the individual level to account for repeated observations on the same monkey. PR and SPR are binary indicators of the respective rearing statuses ($D_{i,PR}$ and $D_{i,SPR}$, respectively).
*$p < 0.05$.