Many child care professionals still believe that orphanages are bad for children. Supposedly, young children fail to develop as they should, socially and psychologically. The prospective harm is even more certain if children are admitted as infants and remain for several years. Supposedly, this bad reputation is based on careful research, not *Oliver Twist* stories of gross maltreatment.

This review seeks to answer two questions: How adequate is the research on orphanage care and are the criticisms of orphanages justified? Several criteria guide the review: 1) Were appropriate questions asked about orphanage care? 2) Were the research methods suited to the questions asked? 3) Do the conclusions correctly follow from the methods used and the data collected? and 4), how widely may we generalize the findings?

Many journals and books, dating back to the 1930s, were searched for relevant material. These were published by professionals in the fields of psychiatry, social work, and psychology. Some worked directly with children but most served as consultants to child care workers or as teachers and researchers. Autobiographical accounts, which can give insights into orphanage care, were not considered pertinent to this review. The term "orphanage" is used by these researchers quite broadly. Sometimes it refers to hospital wards, nurseries, or foundling homes. At other times it refers to boarding homes and other forms of group care. Despite the general lack of detail on how children were treated, we can assume that all these institutions dealt with children who lacked easy access to their parents.

**Perspectives on the Study of Orphanage Effects**

Early in the 20th century, leaders in child care expressed the view that family care is preferable to orphanage care of needy children. This view seemed based on personal preference rather than careful comparison of their relative merits. The available reports were mostly anecdotal or isolated case studies. By 1930, trained social workers helped place children in orphanages or foster care but they seldom evaluated general outcomes. Psychologists had already standardized ways to measure individual intelligence or social maturity but these tools were seldom used outside of schools and clinics. Some clinical psychologists and psychiatrists evaluated selected orphanage children on an individual basis. But this did not constitute evaluations of orphanage programs as such. One important exception is the large scale survey conducted about 1929 (See Trotzkey). Clinical psychologists were used to compare individual intelligence for several thousand orphaned children and home-raised children and they got comparable results. However, the former showed greater gains in body weight than the home-raised children. No study of this magnitude has been conducted since that time!

Psychological theories about child development and ways to handle children had only a slight, indirect bearing on orphanage practice. For example, so-called Behaviorists during the 1920s studied the emotional development of children. Their limited findings encouraged regular, somewhat rigid, eating and sleeping schedules. No systematic studies on the validity of these teachings were conducted in the
orphanage setting. Freudian theories had some impact on child-rearing methods after the 1930s, when child-centered theories became popular. Psychiatrists with Freudian views recommended self-demand schedules, plus liberal weaning and toilet training methods. The aim was to prevent frustration, which might cause neurotic habits. Again, these suggestions were derived mostly from case studies of adult neurotics rather than systematic study of orphanage youth.

About 1940, wartime conditions in Great Britain brought attention to the plight of children separated from their families. Authorities, hoping to reduce casualties from the bombing of large cities, removed children from their families and placed them in special "residential nurseries." These group homes were located in isolated rural communities. Several psychiatrists warned of the risks (Langmeier and Matejcek, 1975; Spitz, 1945). Among these was Anna Freud, who had escaped to England with her father, Sigmund Freud. She was trained in psychoanalysis and committed to the study of child development. Invited by authorities to study the problem of child separation, she concluded from several case studies (Freud and Burlingame, 1944) that institutionalized children are doomed to fail psychologically because of maternal deprivation. This was despite good physical and social care.

This grave warning became widely accepted gospel with the post-war work of John Bowlby, a psychoanalytically trained psychiatrist from the Tavistock Clinic in London. He was commissioned by the World Health Organization to study maternal deprivation and recommend solutions. Limiting himself to European and U.S. findings, he concluded that institutional child care was in a terrible state. In too many cases, the physical care of children was inadequate. Even more serious was the psychological damage which many institutional children suffer despite good physical care. The cause, he assumed, was disruption of the special mother-child bond needed for healthy psychological development. His 1951 report, Maternal Deprivation, was translated into several languages and circulated world-wide. Since his warning served the child care professionals' desire to replace orphanages with home care, the issue of "maternal deprivation" became a central concern for child care workers.

Another psychiatrist, Rene Spitz (1945), provided timely publicity. He reported on the emotional and physical regression of infants in a foundling home, which he believed resulted from maternal separation. His film on the subject, named "Grief" (Time, 1952) further mobilized professional opinion against institutional care. Empirical support came from the clinical studies of William Goldfarb (1943, 1945, 1947, 1949) who tested small groups of adolescents from one New York City orphanage. His reports of serious deficiencies with speech, intellect, personality, and social development were widely circulated among social work professionals. Other scholars in England and the U.S. began to report similar problems. (See the main findings below.)

A bleak view of orphanage care developed within professional child care ranks. It was based more on selected clinical studies than on systematic evaluations or comparisons. In brief, it claimed: -- Any amount of orphanage experience is harmful. The damage is greatest during the first years of life and increases dramatically with length of stay in an institution. Besides being irreversible, the resulting damage affects a wide range of psychological and social traits.-- Supporting popular opinion is illustrated in the extreme views of New York City's Mayor Laguardia, who exclaimed, "The worst mother is better than the best institution." (Goodwin, 1994)

By 1960, child development researchers (Orlansky, 1949; O'Connor, 1956; Langmeier and Matejcik, 1975; Stone, 1954) wondered just what does "maternal deprivation" really mean and is the evidence for its effects valid? New research on child development included both human and animal studies and it partly discounted strong claims about the bad effects of maternal deprivation.

The demise of most orphanages by the 1980s limited further research with orphanage subjects. A few studies continue to be reported outside of England and the United States in third world countries where physical and social facilities are limited. Many of the orphanages that remain in the west have changed to treatment centers for special needs children who are delinquent or emotionally disturbed.
By and large, the child welfare profession and much of the general public stand opposed to any child care option that is not family-centered. For example, Ford and Kroll, who are active in adoption work (1995) state, "Fifty years of research reconfirms the same findings: long-term institutionalization in childhood leads to recurrent problems in interpersonal relationships, a higher rate of personality disorders, and severe parenting difficulties later in life."

**How Crucial Is Maternal Deprivation?**

Since the concept of "maternal deprivation" is central to arguments against orphanage care, it deserves careful study. Its theoretical importance stems, of course, from basic assumptions about the needs of developing children. All of us agree that growing children need strong, interactive relationships with responsible adults. Besides giving emotional and physical security, such ties help the child grow and learn to cope with an ever-changing world. In most societies, the child's parents are considered the optimal social arrangement for child care and mothers are expected to play the central role.

The question arises, what are the risks when a child is separated from its family for whatever reason? Can the mother's role be supplanted to a reasonable degree? Freudian theorists think not. Like Anna Freud (1973), they believe the absence of a close, continuous relationship with a caring mother, or surrogate, spells doom for the psychological well-being of the infant. In layman terms, these theorists argue that "mother-child bonding" is a necessary step toward developing a sense of trust in others, self-confidence, and a sense of right and wrong. This bonding process is assumed to overlay unconscious ego and superego developments, necessary for later psychological health. Such imagery appeals to family-centered professionals who want the authority of scientific theory. Unfortunately, several questions remain unanswered. Precisely what transpires during this bonding process? How well can surrogate-mothers substitute for biological mothers? And, what is the evidence that shows group homes cannot meet this need?

Since the claimed unconscious processes are inferred and never directly observed, their validity remains in doubt. Romantic images of the mother and child looking happily to each other do not themselves discount the possibility that other mothering persons can replace the natural mother. One critic (Orlansky, 1949) notes that most psychoanalytic theories about personality development and its disorders were derived from conversations with adult neurotics. Sears' classic book (1943) reviewed studies designed to test psychoanalytic theory and concluded that no investigator prior to his review actually concentrated on infant behaviors and their relation to personality development. Generally, it is not their style to conduct systematic surveys of large samples in a range of conditions. Rather, they generalize from a small number of selected case studies.

What seems to be an exception to this reliance on isolated clinical cases is Rene Spitz' study (1945; Spitz and Wolf, 1946) of 60 -- some infants in a foundling home. He presents a so-called longitudinal study lasting eight months in which he followed up infants who were admitted soon after birth. Their mothers, who could not care for them independently, nursed them for several weeks or months, then left their care to others. Besides clinical observations of the infants' emotional and physical conditions, Spitz and his staff used the Hetzer-Wolf Baby Tests to monitor their physical and social development while the mothers were present, then absent.

Spitz' graph of average developmental quotients (DQ) shows a dramatic decline which coincides with the departure of mothers. From these data and clinical observations, he concludes that infants separated from their mothers by more than six weeks develop a syndrome of disorders, which he calls "hospitalism." It is marked by tears, staring eyes, and other morbid signs of depression, plus a significant drop in developmental quotient (DQ). Worse yet, 30 percent of these infants died during their first year! Spitz contrasted this image of wasted children with the more positive picture of infants in a prison nursery who thrived on the care of their inmate mothers. This appealing contrast continues to be
A devastating criticism of Spitz' work appeared in a prestigious psychology journal (Pinneau, 1955). One problem is Spitz' confusing description of his infant subjects and the testing procedures. It is not clear how many infants were observed at different stages; what their family backgrounds were; what their conditions upon admission were like; and so on. Apparently, the foundling home was located in Latin America where cultural attitudes about children of deprived or unwed mothers might differ from North American attitudes in the 1940s.

After piecing together numbers from several different reports, the critic concluded this study was not truly longitudinal. The graph which appears to show a decline in average DQs for the same infants at different months of age is actually based on overlapping groups of babies at different ages. The Hetzer-Wolf Baby Tests for measuring DQs were found by other critics to be poorly standardized. For example, these particular tests continue to give lower DQs with successive months of age. Thus, it is not surprising if Spitz' infants showed a lower DQ at a later age. Even more telling are Pinneau's re-calculations which show that average DQs declined sharply before the mothers departed! These facts discount Spitz' reported drop in DQs with months of separation. Pinneau concludes, despite his personal preference for mother care over institutional care, we do not yet have convincing evidence that mother separation rather than social deprivation causes psychological deterioration.

It has also been argued (Ribble, 1943) that emotional bonding during the first weeks of life is absolutely essential if underlying brain cell connections are to be completed, along with other physiological developments. Failures to bond interrupt these changes and causes serious physical damage. The same critic, (Pinneau, 1950) demolishes much of this particular argument with convincing accounts of the neonate's physical development. He reports, also, that few infants reliably recognize their mothers before three months of age. Other child psychiatrists time emotional bonding at six to nine months, when overt recognition is obvious.

The concept of "maternal deprivation" is often used inconsistently. Yarrow (1961) notes it might be equated with orphanage placement, mother neglect at home, or the change in caregivers within the same home. Such variable usage ignores important qualitative differences and contributes to the tendency to not describe orphanage care in detail. Rheingold's time samples (1960) of mothering behavior in family homes and institutions illustrate the kind of careful observation that is needed. She found that home children got contact with more material objects than did orphanage children but the activity levels of both were equal. Prugh and Harlow (1962) draw attention to patterns of "masked deprivation" within family homes, where parents have dysfunctional ties with their children. Parenting styles also vary radically with different socio-economic groups. An example is the wealthy child who is placed in boarding schools at an early age. Even middle class parents may change parenting styles, perhaps due to full-time employment for both. Long hours in a day care setting might also constitute "maternal deprivation" for some children.

Several authors (Bowlby, 1951; Freud, et al, 1973) view staff turnovers in institutional practice as a serious risk to children, since bonding with a single adult is difficult. This argument is tempting but we lack hard, systematic data. Anthropologists report frequent use of multiple-mothering in other cultures. Margaret Mead (1962), for example, suggests this provides healthy continuity of care and a hedge against separation trauma. Kagan (1984) describes a nomadic tribe in Africa which believes the infant is better off if nursed by someone other than its mother. Later, it is cared for by different persons in the tribe. Of course, we need more clear information about the effectiveness of these different child rearing patterns but they do point up the importance of consistent mothering or social support, not biological mothers, per se.

The Israeli Kibbutz has received more consistent study. There, infant mothering is shared by the actual mother and a paid child care worker. The former gives emotional support while the latter attends to physical and personal education through most of the day. Rabin's studies of kibbutz children (1957,1958) found some slowness in personal and social development for one group compared to
non-kibbutz children. Another study showed that nine to eleven-year old kibbutz children were more mature than the comparison group. Thus, local circumstances rather than universal effects of shared mothering seem to explain development outcomes.

An extreme example of multiple-mothering is the "home management house." These occur in educational settings where women students learn infant care under the supervision of one adult. Although research is limited, one study (Gardner and Swiger, 1958) reported no differences between a sample of babies used in one house with regular home babies on measures of physical growth and the Gesell Scales for behavioral and social development.

We conclude that proponents of mother deprivation theory have an appealing theory but few supporting facts. Although emotional and physical deterioration surely occurred in some of their clinical studies, there is insufficient evidence to rule out physical and social neglect as essential causes. Thus, it is far from proven that simply losing one's mother or being placed in an institution by itself leads to psychological decline. Even if we accept the role of maternal deprivation in the case of tiny infants, we have much less evidence that the same argument applies to older children who are admitted to orphanages.

**Reported Psychological Effects of Orphanage Care**

At this point we will evaluate reports on the psychological development of children with orphanage experience. Tables 1 and 2 list many of the studies cited by child care professionals. For convenience, they are grouped into those published before 1951, the year of Bowlby's major review, and those published later. Several studies which dealt with infants less than one year old are omitted since psychological evaluations at this age are difficult or unreliable. Also, this listing is far from complete but it includes most of those cited by critics of orphanage care.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Design</th>
<th>Sample</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bender</td>
<td>1945</td>
<td>Impressions</td>
<td>100s of problem children referrals to hospital</td>
<td>Recalled 5-10% had institutional background</td>
</tr>
<tr>
<td>Bodman and McKinlay</td>
<td>1950</td>
<td>Direct Correlation</td>
<td>51 young adults with orphanage backgrounds, 52 from families, same age</td>
<td>Some orphans found less sociable or mature, poor work adjustment</td>
</tr>
<tr>
<td>Bowlby</td>
<td>1944</td>
<td>Retrospective</td>
<td>44 teenage thieves</td>
<td>A large proportion had orphanage backgrounds</td>
</tr>
<tr>
<td>Bowlby, et al.</td>
<td>1956</td>
<td>Follow-up</td>
<td>60 7-14 year olds hospitalized as infants, 180 classmates</td>
<td>The previously hospitalized differed some in IQ, behavior adjustments</td>
</tr>
<tr>
<td>Author</td>
<td>Year</td>
<td>Design</td>
<td>Sample</td>
<td>Results</td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
<td>-----------------</td>
<td>------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Goldfarb</td>
<td>1943</td>
<td>Direct Correlation</td>
<td>15 12 year olds, orphaned, and fostered; 15 fostered only</td>
<td>Orphaned group lower on personality and intelligence tests</td>
</tr>
<tr>
<td>Goldfarb</td>
<td>1944</td>
<td>Direct Correlation</td>
<td>15 12 year olds, orphaned, and fostered; 15 fostered only</td>
<td>Less emotional-perceptual maturity in responses to ink-blots for orphans</td>
</tr>
<tr>
<td>Goldfarb</td>
<td>1945</td>
<td>9-mos follow-up</td>
<td>15 orphans under 4 years; 15 fostered under 4 years</td>
<td>Orphans rated lower on a wide range of personality and ability tests</td>
</tr>
<tr>
<td>Goldfarb</td>
<td>1947</td>
<td>Retrospective Study</td>
<td>15 well- and 15 poorly adjusted with orphanage experience</td>
<td>Poorly adjusted were placed in home sooner and stayed longer</td>
</tr>
<tr>
<td>Goldfarb</td>
<td>1949</td>
<td>Direct Correlation</td>
<td>15 orphans, 15 fosters, 15 schizoid: all 12 years</td>
<td>Ink-blot responses show immaturity for orphans and schizoid</td>
</tr>
<tr>
<td>Lowrey</td>
<td>1940</td>
<td>Case Study</td>
<td>28 boarding home children with 3 years in orphanage</td>
<td>Showed isolation-type personality (Self-centered, hostile, insensitive)</td>
</tr>
<tr>
<td>Skeels and Dye</td>
<td>1939</td>
<td>Follow-up</td>
<td>13 selected orphan infants; Remaining infants stayed</td>
<td>13 placed with feebleminded women gained in IQ; the others decreased</td>
</tr>
<tr>
<td>Skeels, et al.</td>
<td>1938</td>
<td>Controlled Experiment</td>
<td>Matched pairs of orphan children put into E and C group</td>
<td>Experimentals who got pre-school training gained dramatically in IQ</td>
</tr>
<tr>
<td>Spitz</td>
<td>1945</td>
<td>8-mos follow-up</td>
<td>69 infants, prison nursery; 61 infants, foundling home</td>
<td>Dev. Quotients increased for prison and decreased for foundling babies</td>
</tr>
<tr>
<td>Trotzkey</td>
<td>1930</td>
<td>Developmental and Correlation</td>
<td>2,523 children in orphanage care; 1,214 foster care; NYC school</td>
<td>Physical health and IQs equal or better than Chicago or NYC school norms</td>
</tr>
</tbody>
</table>

Table 2. Reported Psychological Effects of Orphanage Experience: Studies Since 1951
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Design</th>
<th>Sample</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provence</td>
<td>1989</td>
<td>8 mos. follow-up</td>
<td>75 1-yr olds from orphanage, 75 1-yr olds fostered</td>
<td>Orphaned slower to develop social, language, and intellectual skills</td>
</tr>
<tr>
<td>Quinton, Rutter and Liddle</td>
<td>1984</td>
<td>14-yr follow-up</td>
<td>81 female orphan adults, 51 matched controls</td>
<td>Orphaned showed more problems with social interaction, parenting</td>
</tr>
<tr>
<td>Rutter, Quinton and Hill</td>
<td>1990</td>
<td>About 10-yr follow-up</td>
<td>120 with orphanage experience, 58 same age controls</td>
<td>Ratings show more personality and behavioral problems for some orphanage experienced. Associated with parents' problems</td>
</tr>
<tr>
<td>Tizard</td>
<td>1979</td>
<td>6-yr follow-up</td>
<td>20 orphanaged, then adopted, 10 orphanaged, returned parents</td>
<td>Adopted adjusted better than the returned; latter had family problems</td>
</tr>
<tr>
<td>Tizard and Hodges</td>
<td>1978</td>
<td>3.5 yr. follow-up</td>
<td>51 orph (25 later adopted, 13 returned); 29 controls</td>
<td>Teachers saw more problems than did parents with orphaned children</td>
</tr>
<tr>
<td>Tizard and Joseph</td>
<td>1970</td>
<td>Direct Correlation</td>
<td>30 nursery 2-yr olds and 30 2-yr olds from city</td>
<td>Nursery kids less well developed language and mental skills, more fearful</td>
</tr>
<tr>
<td>Tizard and Rees</td>
<td>1975</td>
<td>Direct Correlation</td>
<td>All with orphanage background, age 4.5; 26 stay, 24 adopted, and 15 returned to family</td>
<td>The orphanage remainers had more problems than adopted and less attached to adults. All had same IQ=100</td>
</tr>
<tr>
<td>Wolkind and Rutter</td>
<td>1973</td>
<td>Retrospective Study</td>
<td>All 10-11 yrs old, 458 with behavioral problems, 100 with psychiatric problems, 172 controls</td>
<td>More behavioral problem youth had orphanage background plus prior family disturbances</td>
</tr>
</tbody>
</table>

The studies within each table are alphabetically arranged by author. The underlying research design and sample descriptions are also reported, along with major results. Most of the designs are direct correlation studies, wherein measures such as IQ-scores are compared for a group of orphans and a group of family-raised children, perhaps matched by age and gender. Several designs amount to impressions, case studies, or retrospective reports. The several "follow-up" studies listed here should not be confused with genuine longitudinal studies that measure change over time. These follow-ups simply identify orphans at one stage of development and evaluate them at a later point in time. A small number of these studies are true experiments wherein some treatment variable is directly manipulated by the investigator. Here, subjects are randomly assigned to treatment and control group to insure greater equivalence.

It seems necessary to point out that while much of this research is designed to evaluate the psychological effects of orphanage experience, almost none of it deals with orphanage experience directly. Thus, what ought to be the primary independent or causal variable, whose effects we are concerned about, is ignored until it is time to state conclusions. In many cases orphanage experience is assumed, without evidence, to be sterile or unchallenging. In contrast, home care is assumed to be universally warm and stimulating. Unlike Bowlby (1951), who forgave procedural faults when reviewing many of these same reports, we will consider the effect which these slip-ups have on the scientific credibility of results.
How Do Orphans Compare With Other Children?

Almost all the research on orphanage children amounts to clinical evaluations of opportunistic samples. Since the method of selecting these children is seldom explained, we are unsure how representative they are of a particular orphanage or orphanage children in general. Typical studies assess one or two traits such as intelligence, personality, language skills, or social maturity, usually with the help of trained examiners who use standardized tests.

The most frequently posed question in these studies is -- How does the psychological development of children with orphanage experience compare with other children? The comparisons might be made with general population norms, selected foster children, or children raised by relatives. A glance at the results in Tables 1. and 2. shows the vast majority of these are negative. That is, when groups of children with orphanage experience are compared with other children, the average development for orphans is lower on measures of general intelligence, personality, language, or social skills. There are wide individual differences within some groups of children, but the overall trend is negative for orphaned subjects.

The research methods used partly explain this negative trend. For example, some of the findings amount to impressions or simple anecdotes, not objective evidence. Such reports are prone to reflect personal bias. Bender (1945) recalls how 10 to 20 percent of the problem children referred to the Bellevue Hospital in New York had orphanage backgrounds. This impression might have prompted further research to verify her suspicions but we are not told of that. Neither are we told of base rate norms for admission to this hospital. We might suppose that needy children from socially and economically depressed homes are the rule. Thus, it is no surprise if a portion of these children had been placed in orphanages. Verry's anecdotal report (1939) describes one single graduate from an orphanage school who was unprepared to handle money, job demands, or the requirements for dating girls. Had she never met an immature youth from normal homes? Although she suggests that sheltered conditions in the orphanage caused this unpreparedness, she gives no hard facts. These two reports are frequently cited in child care journals. We need more systematic data to draw serious conclusions about the social maturity of orphanage graduates.

Intelligence or related cognitive skills are widely reported to suffer from orphanage experience. One of the most substantial studies (Trotzkey, 1930) to discount this claim is seldom mentioned, perhaps because it was published privately. It compared orphanage and foster care children in Jewish-sponsored programs in Chicago, New York, and Cleveland with public school children. Over 3,700 Jewish care children were involved, which included 2,523 from orphanages and 1,214 in foster homes. Data on IQ scores, using either group tests or the individually administered Stanford-Binet, showed the Chicago orphanage samples scored in the average range, the same as children in New York City elementary schools. The New York City orphanage samples were slightly lower than the comparison groups but well in the average range. We might add, the borderline or retarded child was not eliminated from these orphanage samples as was usually the case in public schools. The slight advantage which orphanage children had over foster children was suggested to be the result of better education programs for the orphanage children. One part of the study showed that orphanage children, despite being somewhat more underweight upon admission, gained more weight than foster care children. Both groups averaged slightly better gains than public school children during the study period.

Goldfarb's work with small samples from a single New York City orphanage are widely cited by orphanage critics. One study (1943) included IQ test results with tests for speech, educational achievement, and personality. He compared fifteen 12-year old orphans, later moved into foster care, with fifteen foster care-only youth. The former scored well below average on a standardized intelligence test while the latter scored close to average, or about 100. Only three of the fifteen orphans showed average language skills while most of the foster youth did. Educational achievement, ratings for social maturity, and other measures showed the orphan sample to be immature and underachieving.

A serious problem with this finding is the uncertain method of selecting the orphans for study. This,
with the small samples sizes, raises doubts about how representative his subjects were of New York City Jewish orphans. Trotzkey's results, which showed normal IQs, were based on 2,212 orphanage and foster care children in that same city.

Bowlby et al (1956) studied former hospital patients who were believed to suffer maternal deprivation during their one to two year stay in a tuberculosis sanitarium before age four. They compared the 60 former patients, now age 7 to 14, with schoolmates on measures of intelligence, ratings for school behavior, and rated performance while taking the intelligence tests. Average Stanford-Binet IQ's for the two groups were found not to differ statistically. However, a look at just part of the data showed that patients admitted before two years of age more often scored in the lower average range than patients admitted between two and four years. The hypothesis that the former patients would score lower then got indirect support.

Somewhat questionable is Bowlby's treatment of the teacher ratings for school behavior. They found no significant differences in the average ratings for the hospitalized and control groups on each of the 28 rating items. However, the investigators searched through the teacher ratings and threw out those data judged to show inconsistencies. That is, if a teacher rated a given student as "friendly" on one item and "has few friends" on another, his data along with those for his matched controls were eliminated. For the much reduced sample, 11 of the 28 rating items then showed significant differences in the desired direction. Besides other questionable manipulations, the authors equivocated about another finding that displeased them. The teachers happened to rate the former patients as adequate with making close friendships. Having predicted low ability to make friends, the authors conclude that either their rating items had been too crude to show deficient social skills or the subjects had covered up by making superficial friendships. This is a damned if you do and damned if you don't interpretation. The authors' discussion of results suggests an over-zealous attempt to fault the previously hospitalized sample, who were long ago restored to their families.

Several investigators (See Goldfarb, 1943, above) report problems with language skills. It is usually assumed, without direct evidence, that the orphanage offers inadequate adult models for speech. Neither is there evidence to show what is the minimally sufficient speech contact with adults. Evidence of speech development in these studies came from standardized tests or from informal observations. Pringle and Bossio (1960) evaluated a teenage sample from England which had orphanage experience and found below average language and reading skills for those rated "stable" or "maladjusted" by teachers. Since most were currently in foster care, we need more information about the social status of their host families to rule out explanations other than orphanage background. Some investigators (Goldfarb, 1943; Provence, 1989) report lower language development scores and impressions of poor language during testing for other skills. While the testing and observation procedures seem adequate, we seldom learn if the orphanage sample under study comes from socio-economic groups with characteristically poor language skills.

Studies of personality used standardized self-reports, ratings by classroom teachers, and subjectively interpreted inkblot responses. Brown's larger scale study (1937) tested over 200 teenage boys and girls from one Philadelphia orphanage. He compared "Neuroticism" scores for these orphans with the scores for a sample of boys from poor homes and a sample of youth from more advantaged city homes. The orphanage subjects, especially girls, were more neurotic than the city youth but they differed little from boys in lower class homes. He suggested the orphan youth might have come from a lower economic class.

Two studies by Goldfarb (1944 and 1949) used Rorschach ink-blots to assess personality. Comparing 15 teenagers in foster care who had orphanage backgrounds with 15 foster care-only subjects, he found the former were much less mature emotionally. They were also not much different from a small sample of schizophrenic children. He concludes that foster care for these children did not help them overcome their orphanage experience. Without better evidence of the foster care conditions and their personality before admission to the orphanage, we find this conclusion tenuous. He gives some details on the socio-economic status of the foster care families, but there is no explanation of how the samples were
chosen or how representative they were.

Goldfarb also notes there are risks with using subjective interpretations of inkblot responses. Many clinicians, we might note, prefer to use such inkblots only for impressionistic not psychometric comparisons. This is partly due to the unreliability of percentage data for numeric comparisons. Except for reporting the use of independent test examiners in his work, Goldfarb does not convince readers that he coped well with his own personal bias.

Several studies focus on social adjustment and related behaviors. For example, Bowlby (1944) found that several individuals in his sample of teenage thieves had orphanage experience in their backgrounds. This retrospective study, which imputes a causal connection between orphanage experience and criminal behavior, fails badly to rule out alternative explanations. Thus, the conclusion is poorly justified by the evidence. We wonder how Bowlby came to choose this sample? What unreported details about family conditions might explain either their orphanage placement or their criminal behavior?

Other studies of social and behavioral adjustment of orphans offer more details to explain background influences. One (Bodman et al, 1950) compared occupational adjustment and other behaviors for 51 young adults and 52 controls, all from England. During the first year of work, 14% of the orphans failed on the job and returned to the orphanage, compared with none of the controls. Fewer of the orphans participated in social clubs or had boy and girl friends. A 14-year follow-up (Quinton et al, 1984) of young adult females found the orphan subjects had fewer social and parenting skills than matched controls. Tizard's 6-year follow-up (1979) of younger orphans found they adjusted better socially than others from the same institutions who were returned to their families. This atypical trend was explained by the latter group's home situation, where family conflict prevailed. Rutter et al (1990) also show some excess of personality disorders and criminal behaviors in young adults with orphanage backgrounds. As youths, they were placed there because of parental problems with drugs, alcohol, and mental illness. These authors stress that some of these children turned out satisfactorily and that later life conditions were correlated with these personality or behavior problems.

The procedures for these more recent studies on social adjustment seem fairly adequate. Without definite knowledge of the validity and reliability of the instruments used, we have to accept the results at face value. What is seriously missing in these studies is an adequate accounting of the sample selection methods. The resulting inability to know which population groups are actually represented weakens our confidence in the the findings.

Does Early Placement Affect Development More?

A second research question asks -- What are the effects of early placement or prolonged stay in orphanages? The available data on these two topics is extremely sparse. A retrospective study by Goldfarb (1947) compares two groups of 15 teenagers each, both with orphanage experience. One was rated as "poorly adjusted" and the other "well adjusted" by classroom teachers. The former group had more individuals placed in the nearby orphanage before 6 months of age; the latter group was placed closer to one year of age. The poorly adjusted also stayed at the orphanage about three years compared to two years for the well adjusted.

This retrospective study, which suggests a causal connection, gives very weak support for this explanation. It lacks controls for alternative causes such as early family discord which might have led both to orphanage placement and behavior disorder. Such correlation data are best used to design future studies which carefully test for causal relationships.

In the absence of systematic data on the effects of age at first placement and on the effects of prolonged stay in orphanages, it is premature to draw serious conclusions. Clearly, we need far more careful comparisons of children admitted to different institutions at different ages, plus data on length of stay and the interaction of age at placement and length of stay. Also related these two issues is the matter of sibling support. Keeping brothers and sisters together might ameliorate several adjustment problems.
Can Special "Mothering" or Social Stimulation Improve Individual Development?

This research question considers ways to accelerate individual development beyond the expected rate. It suggests the need for a controlled experiment which would provide the proposed stimulation to some individuals and not to others.

Some investigators report quasi-experiments which seem to test hypotheses about the treatment variable. For example, a widely quoted study by Skeels and Dye (1939) described how 13 infants were moved from an orphanage to a home for feeble-minded women. The infants gained dramatically in measured intelligence compared with those infants who remained in the orphanage. The latter group's intelligence actually declined.

The previously cited study by Spitz (1945) also compared infants raised by their prison mothers with infants in a foundling home, where direct mothering was inconsistent. The prison babies thrived under the spoiled attention they received, compared with the deprived babies. This difference was measured by their Developmental Quotients. Spitz realistically notes that the prison babies would likely suffer in later years from the undisciplined, spoiled attention. Neither study, just above, was a true experiment since the investigator did not assign infants to treatment groups nor control the amount of "mothering" received. This leaves us in doubt about the precise role of "mothering", which varies in extent and quality.

Finally, the Iowa Child Welfare group reported several studies in the 1930s which asked -- Can a special pre-school program increase the intelligence quotient for orphanage children?

Wellman and Pegham (1944) report on one of these studies. A controlled experiment gave pre-school training to one group while the matched control group remained in a less stimulating cottage program. The first group's mean IQ increased from below average to average; the other group's below average IQ declined even further. Another report on the same series of studies (Skeels, et al, 1938) describes the living quarters of this orphanage in some detail. It is a an extreme picture of rigid behavior controls and cramped space, with few amenities. Such conditions are not fairly ascribed to all orphanages?

A national expert in research design and psychometrics (McNemar, 1940) takes Wellman and others in the Iowa Child Welfare group to task. He charged them with mishandling statistical analyses and with possible bias when testing children. Generally, the psychometric profession is skeptical of raising IQs significantly (e.g., a full standard deviation) except in the case of acutely deprived children who might be brought up to their potential. This latter point could explain the Iowa group's reported success and it further testifies to the extreme degree of repression in that particular orphanage.

General Criticisms

There is no denying the negative picture portrayed by many of these published reports. Should we accept them at face value and assume it is a scientific fact, orphanages are not good for children? Before answering, we should consider several points.

Were the Research Questions Appropriate?

Most of this research centers on a single issue -- How do orphan children compare with other children psychologically? While this is a legitimate outcome question, it does not deal directly with orphanage experience, or the root causes of psychological development within the institution. Almost never are direct, systematic observations made of the quality of orphanage care or the kinds of social experience provided. We learn nothing about the sleeping, eating, and moral habits of orphanage children. Are these habits different, better, or worse because of the regular schedules and predictable rules that characterize
many institutions?

The orphanage setting offered potentially rich opportunities to study child development but this was rarely considered. Other questions might have focused on the advantages or disadvantage of large peer groups, associated with the segregation of similar age children and genders within one group cottage. What are the positive or negative effects of living with large peer groups? How did the required work, to make beds, clean cottages, mow lawns, prepare food, and so on, affect individual character or personal habits?

Assessments of outcome could be broadened to include surveys of achievement after leaving the orphanage. Two unpublished studies illustrate this point. One reported that the World War II non-orphaned draftees in North Carolina showed a rejection rate of about 20 percent for physical fitness compared with 2 percent for draftees from orphanages in that state. A second study (McCall, 1951) reported that the occupational achievements for graduates from one particular large orphanage followed the same distribution as census norms for occupational levels in the U.S. That is, there were equal proportions of professionals, skilled workers, unskilled, and so on.

Were the Research Designs Appropriate?

Several professional journals in the child care field report anecdotal or retrospective findings which are frequently cited by others as evidence of psychological effects from orphanage care. Such data might be justified as background work, aiming toward hypothesis development, but they should not be taken as conclusive evidence.

This field of research leans too much on correlational designs to test hypotheses about causal relationships. For example, questions about IQ differences between orphans and other children are answered by comparing two groups of children at one point in time. Insufficient attention is paid to background differences other than orphanage experience. The results might give slight support to causal predictions but they never give satisfactory proof. There is always the possibility that some unidentified cause explains score differences. For example, if more orphanage children come from lower socioeconomic classes where intelligence scores are lower, we might expect lower IQ scores for them. (We must not overlook the wide range of differences within each group and their overlap.)

Some of the most polished work was in the form of controlled experiments (Hakimi-Manesh, 1984; Rheingold, 1956) which demonstrated that enriched social stimulation can enhance social responsiveness in infants. Much more work of this rigor is needed to test causal hypotheses about features of child care work.

There is too much reliance on extremely small samples. This is especially risky in correlational designs, where numeric results are sensitive to uncontrolled or hidden factors. Since most psychometric instruments are subject to errors of measurement, large samples are needed to insure more valid comparisons. Curiously, the single large-scale study of IQ scores in this survey of the literature (Trotzkey, 1930) gave a favorable impression of orphanage care.

Finally, the subject matter demands true developmental research. This is the method of measuring changes for the same children over time. Almost none of the research reviewed is developmental although much of it was described as "follow-up" research. It simply identified children at one stage, then evaluated them at a later stage. No actual changes were recorded. Again, the exception was Trotzkey's report of physical weight gains for several thousand children from Jewish orphanages. His results showed that children were underweight for their age when first admitted. Later, the majority were average or better for their age group. Other questions, such as the optimal age for placing children in orphanage care could be much better answered by systematic developmental studies based on samples involving many institutions.

Do Conclusions Properly Follow From the Procedures Used?
Oftentimes, the conclusions did not seem justified by the procedures. The primary reason for these doubts about internal validity was faulty sample selection. Convenience or opportunistic samples were the rule. Almost never were the chosen samples shown to be representative of some given orphanage group. Needless to say, they were not shown to be representative of orphans in general. Yet, very broad generalizations were made about orphanage care.

An example of this problem is the post-war evaluation of young adults who were separated from their parents to escape bombing. Maas (1963) described a sample that adjusted rather well but Bowlby (1951) cites a study which shows one group suffered psychiatric problems. Neither study makes a good case for representative samples.

Another complication is the fact that the subjects for study came from widely different institutions. Some were hospital patients, some were infants in foundling homes, and some were in foster care. This is in addition to the more conventional institutions where children are cared for in large groups. All of these children might be separated from their own families but that is about all they shared.

There was a general tendency to conclude that orphanage care was the cause of reported psychological defects. Yet, mere membership in the orphanage population was the only way that orphanage care was indicated. Actual treatment conditions were often assumed but seldom observed. The majority of such studies used no control for length of stay or other descriptions to show the quality or amount of care received.

The credibility of certain results was diminished by claims that cross-age samples gave developmental data. At best, the design reported by Spitz (1945) is an obscure mixture of these. There also appears to be some lack of restraint in twisting results to make them suit predictions. This is obvious in Bowlby et al's follow-up study (1956) of previous hospital patients where some teachers' ratings were omitted that appeared "unreliable". Other studies had skimpy descriptions of sample selection and testing procedures and failed to show how investigator bias was controlled.

**Why So Many Negative Outcomes?**

We might accept some of this research at face value. Where there is obvious failure to provide for the emotional or social needs of children, we should expect psychological deficits. This concession is far from granting that neglect was a common or general practice.

It seems fair to say that bias against the orphanage care option dominates the child care profession. Thus, we are not surprised if most reports are "politically correct." We might also expect editors to be a bit more lenient with borderline research designs or methods. To some extent, we are dealing with so-called "applied" research which is characteristically less rigorous about rules of procedure. This point is illustrated by Bowlby's frank recognition (1951) that the evidence is largely clinical, not systematic, nor statistically controlled. He concluded that the persistent negative trend had more scientific importance than procedural rules.

Another selection factor is the tendency for editors in most fields to reject studies that show no difference for the comparison groups. For example, it is hard to be sure why a sample of orphans have the same average intelligence as a sample of home care children. There might be no true difference, chance errors might have made it appear so, or mistakes by the investigator might have obscured true differences. Also, since most editors expect orphanage samples to show less intelligence, they are skeptical of exceptions.

**Conclusions**
Several conclusions may be derived from this critical review:

1. Theories about the detrimental effects of maternal deprivation get highly tenuous, indirect support at best from orphanage research. Where psychological deterioration in infants was found, it is not clear whether mother absence or simple physical and social neglect was the essential cause. Neither was there any evidence that such neglect was a widespread practice.
2. Some teenagers and young adults with orphanage experience show deficits in language development, intellect, personality, or social skills. However, it is far from clear that this was caused by their orphanage care. Orphanage care, per se, was almost never directly observed or explicitly manipulated in this research.
3. Most of this research suffers from the overuse of small, opportunistic samples and there is a general failure to describe population sources and methods of selection. These limitations make it impossible to generalize findings based on isolated samples to all orphans or orphanages.
4. Most orphanage research is limited to a narrowly focused, clinical search for psychological damage. Very little of it deals with the effects of age at placement. None of it deals with the role of sibling support, the effects of age or gender groupings, the role of work, moral training, and a host of other practical issues in orphanage care.
5. Critics of orphanage care seem over-zealous to produce negative evidence, then generalize their findings to all orphans or orphanages. More consideration should be given to positive orphanage experiences and ways of assessing their effects. Besides the controlled experiments with infants using social stimulation, there was only one developmental study which directly measured change. More developmental research is needed.

References


Year of Life. NY: John Wiley and Sons.


