
A SOCIOEMOTIONAL INTERVENTION IN A LATIN AMERICAN ORPHANAGE

ROBERT B. McCALL, CHRISTINA J. GROARK, AND LARRY FISH

University of Pittsburgh

DIANE HARKINS

University of California, Davis

GABRIELA SERRANO AND KAREN GORDON

Whole Child International

ABSTRACT: A pilot intervention that emphasized training and technical assistance to promote warm, sensitive, and responsive one-on-one caregiver–child interactions primarily during feeding and bathing/changing was implemented using regular staff in a depressed orphanage for children birth to approximately 8 years of age in Latin America. Despite a variety of unanticipated irregularities in the implementation of the intervention, many beyond the researchers’ control, ward environments improved; caregivers displayed more warm, sensitive, and responsive interactions with children; and children improved an average of 13.5 developmental quotient (DQ) points after 4+ months’ exposure to the completed intervention. Furthermore, 82% of the children had DQs greater than 70 before the intervention, but only 27.8% did so afterward. Although the training for all caregivers was aimed at children birth to 3 years, the number of different caregivers was reduced, and technical assistance was provided only to caregivers serving children less than 3 years, younger and older children (3–8 years) improved approximately the same amount. However, children who were transitioned from a younger to an older ward during the intervention improved less than did children who remained in either a younger or an older ward, the first evidence suggesting that the common orphanage practice of periodically graduating children from one homogeneous age group to another may impede their development. The study is consistent with others that have shown that orphanages can be changed, and increases primarily in warm, sensitive, responsive caregiver–child interactions can produce improvements in children’s development.

The intervention and evaluation reported here was funded by Whole Child International, Karen Gordon, Founder and Executive Director. The intervention was created and implemented by Whole Child International with the help of WestEd and the Pikler Institute of Budapest, and the evaluation was conducted independently by the University of Pittsburgh, Office of Child Development. The authors thank Consuelo Espinosa, Janet Poole, and Edilma Serna for help in training; Auxiliadora Alvarado, Mario Bandes, Ernest Dettore, Ian Downing, Juan Diego Rivas, Raquel Suazo, and Karla Urbina for conducting assessments; Mauricio Gaborit and Marisol Vanegas Guido for local supervision; and Anna Tardos for program consultation.

Direct correspondence to: Robert B. McCall, Office of Child Development, 400 N. Lexington Avenue, University of Pittsburgh, Pittsburgh, PA 15208; e-mail: mcall2@pitt.edu.

INFANT MENTAL HEALTH JOURNAL, Vol. 31(5), 521–542 (2010)

© 2010 Michigan Association for Infant Mental Health

View this article online at wileyonlinelibrary.com.

DOI: 10.1002/imhj.20270

RESUMEN: Se implementó una intervención experimental piloto que enfatizaba entrenamiento y asistencia técnica con el fin de promover interacciones de uno a uno cálidas, sensibles y receptivas entre el niño y quien le presta el cuidado, usando el personal de servicio regular en un deprimente orfanato para niños recién nacidos hasta aproximadamente los 8 años de edad en América Latina. A pesar de la variedad de irregularidades no anticipadas en cuanto a la implementación de la intervención, muchas de ellas fuera del control de los investigadores, el ambiente de los pabellones mejoró; quienes prestaban el cuidado mostraron interacciones más cálidas, sensibles y receptivas con los niños; y los niños mejoraron hasta llegar a un promedio de 13.5 puntos en el cociente de desarrollo después de más de 4 meses de estar expuestos a la intervención completa. Es más, 82% de los niños tenía un cociente de desarrollo de menos de 70 antes de la intervención, pero solamente 27.8% tuvo ese resultado después de la misma. Aunque el entrenamiento para quienes prestaban el cuidado fue dirigido a niños recién nacidos hasta los 3 años de edad, y el número de diferentes prestadores de cuidado se redujo de tal manera que la asistencia técnica se les dio solamente a quienes cuidaban niños de menos de 3 años, tanto los niños más pequeños como los mayores (3-8 años) mejoraron aproximadamente la misma cantidad. Sin embargo, los niños que hicieron la transición de un pabellón para niños pequeños a uno para niños mayores durante la intervención mejoraron menos que los niños que permanecieron en el pabellón para niños pequeños o en el pabellón para niños mayores, lo cual representa la primera evidencia que sugiere que la práctica común en los orfanatos de periódicamente graduar niños de un grupo de edad homogénea y pasarlo a otro grupo pudiera impedir el desarrollo de los niños. El estudio es consistente con otros que muestran que los orfanatos pueden ser cambiados y que los incrementos primariamente en las interacciones cálidas, sensibles y receptivas entre el niño y quien le cuida pueden producir mejoras en el desarrollo de los niños.

RÉSUMÉ: Une intervention pilote ayant mis l'accent sur la formation et l'assistance technique afin de promouvoir des interactions mode de soin-enfant personnalisées chaleureuses, sensibles et réceptives durant les repas et la bain/change ont été mises en place en utilisant un groupe d'employés habituels dans un orphelinat pauvre pour des enfants de la naissance à approximativement 8 ans en Amérique Latine. En dépit d'une variété d'irrégularités non anticipées durant la mise en application de l'intervention, beaucoup d'entre elles n'ayant pas pu être contrôlées par les chercheurs, les services de l'orphelinats ont progressé; les modes de soin ont fait preuve de plus d'interactions chaleureuses, sensibles et réceptives avec les enfants; et les enfants ont progressé d'une moyenne de 13,5 DQ points après quatre mois d'exposition et ce jusqu'à la fin de l'intervention. De plus, 82% des enfants avait des DQ inférieurs à 70 avant l'intervention, mais c'était seulement le cas de 27,8% d'entre eux après. Bien que la formation de tous les modes de soin avait eu pour but les enfants de la naissance à l'âge de trois ans, et bien que le nombre de modes de soin différents ait été réduit, et que l'assistance technique n'ait été donnée qu'aux modes de soin s'occupant des enfants de moins de trois ans, les enfants les plus jeunes et les enfants les plus vieux (3-8 ans) ont progressé à peu près de la même manière. Cependant, les enfants ayant transitionné d'un service pour les plus jeunes vers un service pour les âgés durant l'intervention ont moins progressé que les enfants étant restés soit dans un service pour les plus jeunes soit dans un service pour les plus âgés, les premières preuves suggérant que cette pratique commune dans un orphelinat de faire passer périodiquement les enfants d'une groupe d'âge homogène à un autre freine peut-être le développement. Cette étude s'accorde avec d'autres études qui montrent que les orphelinats peuvent être changés et que les augmentations avant tout dans les interactions mode de soin-enfant chaleureuses, sensibles et réceptives peuvent produire des améliorations dans le développement de l'enfant.

ZUSAMMENFASSUNG: Ein Pilot-Projekt mit Interventionscharakter und dem Ziel, durch Ausbildung und technische Unterstützung eine warme, sensible und angemessene eins-zu-eins Betreuer-Kind Interaktionen v. a. beim Füttern und Baden bzw. An- und Ausziehen zu ermöglichen, wurde innerhalb eines Waisenhauses

für depressive Kindern im Alterspektrum Geburt bis ca. 8 Jahre in Lateinamerika umgesetzt. Trotz einer Vielzahl von unerwarteten Herausforderungen bei der Durchführung der Intervention, die von den Forschern nicht kontrolliert werden konnten, verbesserten sich Umweltfaktoren: Die Bezugspersonen zeigten wärmere, sensiblere und mehr angemessene Interaktionen mit den Kindern; die Kindern verbesserte ihre DQ um durchschnittlich 13,5 Punkten, nachdem sie 4 oder mehr Monate an der Intervention teilnahmen. Außerdem hatte 82% der Kinder eine DQ <70 vor der Intervention, aber nur mehr 27,8% danach. Obwohl die Ausbildung für alle Pflegepersonen die mit Kindern im Alterbereich Geburt bis 3 Jahren ausgerichtet war und die Anzahl der verschiedenen Bezugspersonen reduziert wurde und technische Unterstützung nur für Pflegepersonen des o. g. Altersbereichs angeboten wurde, profitierten sowohl jüngere (0-3 Jahre) als auch ältere Kinder (3-8 Jahre) im gleichen Maße von der Intervention. Allerdings verbesserten sich die Kinder weniger, die von einer Station für jüngere Kinder auf eine Station für ältere Kinder während der Intervention wechselten. Dies weist darauf hin, dass die gängige Praxis der altershomogenen Altersgruppen in Waisenhäusern möglicherweise deren Entwicklung behindert. Die Studie steht im Einklang mit anderen Untersuchungen, die auch aufzeigen, dass das Milieu in Waisenhäusern geändert werden kann und warme, sensible und angemessene Bezugsperson-Kind-Interaktionen zur Verbesserungen der Entwicklung von Kindern beitragen.

抄録：主として授乳と入浴／おむつ交換の間の、暖かく感受性に富み応答的な、1対1の養育者と子ども相互交流を促進するための、トレーニングと技術的な援助を強調するパイロット的な介入が、ラテンアメリカの0歳からだいたい8歳くらいまでの子どものための最下層の児童養護施設の正規職員を用いて実施された。介入の実施における様々な予測していなかった不規則性、その多くは研究者のコントロールを越えた物であったが、にもかかわらず、施設 wardの環境は改善した。養育者はより多くの暖かく感受性に富み応答的な子どもとの相互交流を表した。そして子どもは4ヶ月間の完結した介入を受けた後に、平均13.5ポイントDQが改善した。さらに、介入前は82%の子どもがDQ70以下だったが、介入後は27.8%だけになった。すべての養育者へのトレーニングが0歳から3歳の子どものみを対象にしており、異なる養育者の数は減少し、3歳以下の子どもの世話をする養育者にもみ技術的な援助が提供されたとはいえ、幼い子どもと年長(3-8歳)の子どもは、ほぼ同じ量の改善をみた。しかしながら、介入の間に幼い子どもの施設から年長の子どもの施設に移行した子どもは、それぞれの施設にとどまった子どもよりも改善が少なかった。これは、養護施設の一般的なやり方である、定期的に子どもたちが一つの均質な年齢集団を卒業して別の年齢集団に入ることが、子どもたちの発達を妨げる恐れがあることを示唆する最初の根拠だった。この研究は、養護施設は変えられることができ、主として暖かく感受性があり応答的な養育者と子ども相互交流を増大させることにより、子どもの発達の改善がもたらせられ得ることを示すそのほかの研究と、一致する。

* * *

An estimated 143-million children live without permanent parents in 93 countries in sub-Saharan Africa, Asia, Latin America, and the Caribbean (UNAIDS, UNICEF, USAID, 2004), and UNICEF estimated an additional 1.5 million in Central and Eastern Europe (Nelson, 2007).

These include “true orphans” and “social orphans” whose parents are unwilling or unable to care for them. The latter are given up because of war, natural disasters, and epidemics (e.g., HIV); economic insufficiency; inability of the parents to rear the child (e.g., teenage mothers, drug and alcohol abuse, mental health issues); cultural aversion to children with disabilities; and involuntary removal of the child associated with child abuse and neglect (St. Petersburg–USA Orphanage Research Team, 2008).

Although some children are reared in refugee camps and older children may be living on the street, many of these children are reared in institutions. Most countries at some point in their histories have institutionalized children without permanent parents (Nelson, 2007), and this practice persists in many countries such as the former Soviet States and other countries of Eurasia. It was estimated that 5 to 20 children per 1,000 are in institutions in Eastern Europe (Nelson, 2007).

A few institutions can be quite good (Gavrin & Sacks, 1963), but generally, institutions do not provide an adequate rearing environment (e.g., Rosas & McCall, 2009). Some may be severely and globally depriving (e.g., Romanian orphanages in the 1990s) while others can be acceptable on many characteristics (Gavrin & Sacks, 1963; St. Petersburg–USA Orphanage Research Team, 2005, 2008). Across those institutions described in the literature, however, nearly all provide inadequate socioemotional environments in which there are many, changing, and inconsistent caregivers who deliver impersonal, perfunctory caregiving with little talking, sensitivity, or responsiveness to children (Rosas & McCall, 2009). The result seems to be developmental delay in every aspect of children’s development, including physical growth, cognitive and language development, attachment, socioemotional maturity, and problem behavior (Rosas & McCall, 2009).

A variety of interventions have been tried in institutions, with the aim of improving children’s development (Rosas & McCall, 2009). Some studies had provided supplementary sensoriperceptual stimulation administered by noninstitutional staff for a few hours per day over a few weeks predominately to infants in their first year of life (Brossard & Decarie, 1971; Casler, 1965; Hakimi-Manesh, Mojdehi, & Tashakkori, 1984; Kim, Shin, & White-Traut, 2003; Sayegh & Dennis, 1965). These interventions had mixed results, and those that produced advantages for the intervention children did so partly by minimizing the decline otherwise typically observed in the comparison children.

A second group of studies had provided additional stimulation through caregiver–child interactions delivered in rather specific circumstances to infants and toddlers (Rheingold, 1956; Saltz, 1973; Smyke, Dumitrescu, & Zeanah, 2002; Sparling, Dragomir, Ramey, & Florescu, 2005; Taneja et al., 2002). These interventions tended to produce more consistent improvements in general development, although the size of the improvements were modest and, again, produced to some extent by preventing declines that were otherwise observed in comparison children. Effects tended not to last after the intervention stopped.

A third set of interventions attempted to change the entire institutional environment all day, every day, using the existing staff of the institution or other services (McCall, Groark, Fish, & The Whole Child International Team, 2008; Nelson et al., 2007; St. Petersburg–USA Orphanage Research Team, 2008; Wolff, Dawit, & Zere, 1995; Wolff, Tesfai, Egasso, & Aradom, 1995; Wolff & Fesseha, 1998, 1999). These interventions attempted to reduce the number of different caregivers, increase their consistency in children’s lives, and promote warm, caring, sensitive, and responsive caregiver–child interaction; they did not change the nature and quality of medical care, nutrition, and sanitation. Collectively, such interventions produced improvements in most domains of children’s development, and the more comprehensive and intensive the intervention, the greater the improvements. Thus, these studies pointed to the socioemotional environment as

being a major corrosive factor in many institutions, and improvements in this domain lead to improvements in most aspects of children's development.

Implementing such comprehensive interventions is challenging because it requires changing longstanding policies and behavioral styles of regular staff, some of whom have worked in the institution in this manner for decades. Institutional directors and staff tend to resist change, perceive any requested change as producing more work for them, and may be initially uncertain whether they will be able to implement the requested changes (St. Petersburg–USA Orphanage Research Team, 2008). Furthermore, implementation of interventions in general has not been studied very extensively (Fixsen, Naoom, Blasé, Friedman, & Wallace, 2005), so there are few guidelines for implementing such changes (Groark & McCall, 2008; St. Petersburg–USA Orphanage Research Team, 2008).

The current study represents a pilot attempt by a nonprofit organization, Whole Child International, whose primary aim is to improve orphanages, to implement an intervention stressing the socioemotional aspects of caregiver–child interaction in a depressed orphanage in Latin America. It joins the small set of other similar studies that have primarily emphasized the socioemotional caregiver–child interaction; medical care, nutrition, and sanitation are not modified.

The conditions in this orphanage were stark. Although each group of 8 to 13 children of relatively homogeneous ages had several rooms in a separate building, the rooms were unadorned, dark (to save electricity), and in ill repair. There were few toys and little furniture except children's beds, tables, and chairs. Caregivers worked 12-hr shifts with 2 to 3 days off, and caregivers were rotated to different wards, sometimes after only 2 to 3 months. At any one time, there was only one or two staff to care for 8 to 13 children, and caregivers spent little time talking or interacting with children.

Thus, the questions were whether a socioemotional intervention could be implemented in an entire orphanage using regular staff and if it would produce developmental benefits for children.

PROCEDURE

Participants

Participants were all caregivers and all children birth to 8 years available at pre-intervention assessments in 2006 and postintervention assessments in 2008 in an orphanage in Latin America. Children were housed in wards of 8 to 13 younger (Birth–12, 12–24, and 24–36 months) and older (3–4, 4–5, 6–8 years) children.

The Intervention

The intervention consisted of training all caregivers, providing some of them with technical assistance (i.e., observation and suggestions to promote use of training information in caregiving practice), and making a few supportive structural changes (e.g., employment patterns, furniture) in some wards.

General principles. Most of the training and other intervention activities were drawn from the WestEd Program for Infant–Toddler Care (Gonzalez-Mena, 2005; Lally, 1990; Lally, Mangione, & Young-Holt, 1992) and the Pikler Institute of Budapest, Hungary (Tardos, 2007). Philosophically, these programs emphasized four general principles that guided the current intervention:

- *Caregiving routines are an important time for interaction between adult and child.* Each child should receive one-on-one attention primarily during the caregiving routines of

feeding and bathing/dressing. Adults should describe to children of all ages the child's experience at the moment and what the caregiver is about to do with them. A predictable sequence of caretaking of individual children is established so that over time, children learn their place in the order of care ("Juan is bathed after Lucida.")

- *Continuity of care is particularly important for children in institutions.* Children need a minimum number of caregivers who interact with them in approximately the same way on a regular basis. A goal was to maintain the same caregivers for the same group of 8 to 13 children, although this was only partly achieved; and part-time volunteers were discouraged from providing direct care to children to minimize the number of different caregivers and increase the consistency of care.
- *Children need freedom of movement to grow and learn.* Children need ample, safe places in which to move and explore, primarily on their own and without adult demonstration, coaching, or encouragement. Nonmobile infants should spend their waking hours in large, enclosed spaces rather than in cribs.
- *Plentiful, safe, and developmentally appropriate materials must be accessible to children during their waking hours.* When not receiving one-on-one caregiving, children need safe, plentiful play materials to engage their interests and attention in uninterrupted time for exploration and play, primarily on their own. Thus, caregivers are to focus attention on children during feeding and bathing/changing to a greater extent than during free play.

Additional training topics included following the child's attention and interest, promoting language development, the use of observation, caregiver self-reflection, attachment, early brain development, and adapting to the child's temperament.

Preliminaries. The intervention was created and supervised by several individuals representing Whole Child International, a nonprofit group devoted to improving caregiving environments of orphanages. The Whole Child International Team plus the director, pedagogue (i.e., person responsible for supervising caregivers), and social worker of the orphanage and a representative from UNICEF in Latin America visited the Pikler Institute in Budapest for 5 days of observation, discussions with Pikler staff, and training in the Pikler approach to orphanage care. The visit provided a common experiential basis to the orphanage administrative staff, concrete examples of how the intervention could operate, and enthusiasm of orphanage leaders for implementing the intervention in their institution.

Training team. Whole Child International collaborated with WestEd's Center for Child and Family Studies to conduct the training. WestEd is the largest U.S. trainer of nonresidential early care and education staff. Two bilingual and bicultural WestEd staff members and Whole Child International staff conducted the training and technical assistance 1 week per month for 12 months between November 2006 and October 2007.

Trainees. Essentially all staff of the entire orphanage were trained, including caregivers, aides (who mainly performed housekeeping duties, but sometimes assisted caregivers with children), infirmary staff, day and night supervisors, the Orphanage Technical Team (director, pedagogue, social worker, psychologist, pediatrician, head nurse, physical therapist), and occasionally

kitchen staff at their request. Staff were divided into three groups, and received the same training on 3 separate days so that all staff had the opportunity to attend training on their day off.

Weekly training and technical-assistance schedule. During any single training and technical-assistance week, Monday was spent with the Orphanage Technical Team to develop an action plan for implementing that week's core principles, assess progress to-date, and review the week's schedule of activities. The three groups were separately trained for 6 hr each on Tuesday, Wednesday, and Thursday. Friday was devoted to technical assistance, consisting of observing caregivers on the wards and providing suggestions and encouragement to the caregivers to implement what they had learned in training.

During the year of training, the trainers perceived that behavioral improvements on the wards lagged behind the training, so some of the later weeks were devoted entirely to technical assistance. In addition, the social worker and the psychologist of the orphanage became mentors to staff, helping them practice the training in caregiving behavior that they received, although this was not always consistently or successfully implemented and occasionally produced conflicting directives between these mentors and formal supervisors.

Training and technical assistance topics. Table 1 presents the training and technical-assistance schedule and topics.

The training emphasized caregiving for children birth to approximately 3 years of age, and technical assistance was given to only caregivers serving children in approximately this age range. Thus, although caregivers for older children received the same training, that training was largely aimed at younger children, and they did not receive technical assistance.

Structural changes in different wards. Although the training followed a common schedule for all caregivers serving children of all ages, other structural changes were implemented in different months for different lengths of time and for different wards (for details, see McCall, Groark, Fish, and The Whole Child International Team, 2008). These "structural changes" consisted of identifying for a group of children 4 caregivers who would work 8-hr shifts approximately 5 days per week; thus, children in that ward would be exposed primarily to those 4 caregivers the entire week, every week. However, only 1 caregiver and 1 aide, who primarily did housekeeping chores, but sometimes assisted in caregiving the children, were available at any one time for 8 to 13 children, which is a rather high child:caregiver ratio for infants and toddlers. Caregivers were concerned about traveling to the orphanage at odd hours, so transportation to and from work was arranged for their safety. In addition, caregivers were given a 20% supplement to their wages paid every 3 months throughout the study to cover the extra time and inconvenience engendered by the intervention.

An aim had been to reduce group size, but a sudden influx of infants and toddlers prevented that and even required setting up an additional ward for young infants. Generally, a schedule of the sequence in which children would be individually fed and bathed was established. Additional furniture and play materials were provided, such as bookcases to hold play materials at children's levels, Pikler-style changing tables, and more comfortable chairs for feeding infants. In addition, younger infants spent daytime hours in a large, enclosed and padded space called a "ring" rather than in their cribs. High chairs, which had previously been used for feeding, were removed to encourage caregivers to feed children on their laps. Additional toys and equipment provided

TABLE 1. *Training and Technical Assistance Schedule and Topics*

November 13–17, 2006

- Get acquainted activities
- Developmental milestones of children ages 0 to 3, including brain development
- Overview of relationship-based, responsive, and respectful caregiving
- Experiential activities focused on respectful routines
- Self-reflection: How our own childhood experiences can influence our caregiving

December 4–8, 2006

- Relationship-based, responsive, and respectful caregiving
- The Program for Infant–Toddler Care responsive process of Watch-Ask-Adapt
- Experiential activities focused on respectful care while dressing children

January 15–19, 2007

- Relationship-based, responsive, and respectful caregiving
- The role of self-reflection in providing responsive care, “professional” love
- Experiential activities focused on respectful care while diapering children

February 12–16, 2007

- Understanding children’s temperament
- Introduction of developmentally appropriate play materials
- Experiential activities focused on respectful care while bathing children

March 12–16, 2007

- Technical assistance only (approximately 15 hr of TA in wards, 5 hr of meetings with technical team)

April 16–20, 2007

- Primary caregiving in institutions
- Continuity of care in institutions
- Experiential activities focusing on responsive caregiving during feeding routine
- Introduction of the planned pilot ward

May 20–23, 2007

- Training with technical team and pilot ward staff only, including discussion of the vision, program philosophy, the new work schedule, and the value of free play
- Technical assistance primarily focused on preparing the pilot ward

June 6–9, 2007

- Technical assistance only with a focus on the implementation of responsive caregiving practices and the adoption of a sequence in the children’s routines; 8 children moved into the pilot ward

June 11–15, 2007

- Cognitive and language development in young children; universal precautions (particularly handwashing); and the sharing of childhood stories, songs, and rhymes

July 8–12, 2007

- Technical assistance only, with a focus on the environment and responsive routines

August 5–9, 2007

- Technical assistance only, with a focus on responsive routines, play materials in the environment, and the role of the aide (colaboradora) in supporting the caregiver (educadoras) during routines

September 2007

- Training and technical assistance by Whole Child only with a focus on observation of wards and written recommendations to the director and technical team; team did not coach individual caregivers

October 15–19, 2007

- Observation of young children
 - Documentation, including the use of photos, keeping children’s diaries, and note-taking
 - Adapting the environment and play materials
 - The stages of play in young children
 - Closing celebration
-

included stacking cups, bowls, dolls, small plastic balls, stuffed animals, and a “pañuelitos,” soft flannel cloths for each child with their name on it which became an infant’s first personal toy.

Four- and 5-year-old children had gone to an offsite preschool, but this service was halted when an orphanage child wandered away from the site and was temporarily lost. Thus, older children received no special services during the pre-intervention period. Later, a preschool program was offered within the orphanage, an offsite preschool was reestablished for older children, and swimming and soccer as well as arts activities were arranged for the oldest children before postintervention testing.

Assessments

Assessments were made of the overall ward environment, caregiver–child interactions, and the general behavioral development of children before any interventions were introduced (“pre-intervention”) and at least 4 months after interventions had been completed (“postintervention”), although the staggered implementation of the intervention across wards meant that some wards had longer exposures to the intervention than did others.

Infant/Toddler Environment Rating Scale (ITERS; Harms, Cryer, & Clifford, 2006) and the Early Childhood Environmental Rating Scale (ECERS; Harms, Clifford, & Cryer, 2005). Wards with children birth to 36 months of age were rated with the ITERS; wards of children ranging from 18 to 72 months (although the individual wards were more age homogeneous) were assessed with the ECERS. Generally, 1 caregiver and 1 aide for groups of 8 to 13 children were present during the assessment. These assessments required approximately 2.5 hr of observation and questioning of caregivers, and yielded a Total Score and seven subscale scores ranging from 1 to 7 (best).

A single coder was trained over 2 days by a highly experienced ITERS/ECERS trainer. The training included both trainer and coder together scoring early childhood care and education rooms in the United States and discussing their scores. The single coder did all the pre-intervention scoring. This coder had left the project by postintervention, so the original U.S. trainer was sent to Latin America at postintervention to train another coder; both independently scored each of the seven wards. Although only the trainer’s scores are used in this report, the coders correlated .74 on the seven subscale scores across the seven wards.

Caregiver–Child Social/Emotional/Relationship Rating Scale (CCSERRS; McCall, Groark, & Fish, 2010). A new scale, the CCSERRS, was used to assess the nature of the social and emotional caregiver–child interactions. The scale was created especially to rate the warmth, sensitivity, responsivity, and mutuality of caregiver–child interactions in orphanage contexts. It consists of 18 items to be rated 0 (*never*), 1 (*sometimes*), 2 (*most of the time*), or 3 (*always*). In this study, coders were to rate each caregiver (but not aides) on two 5-min observations focusing on different children for each of three activities: feeding, bathing/changing, and free play. Thus, caregivers were observed for 30 min and given six ratings, which were averaged across the two episodes within each activity.

During the scale’s development, intercoder reliabilities (McCall et al., 2010) were high even for a single, 5-min observation of free play, and validity was demonstrated with substantial differences in scores between an orphanage in St. Petersburg (Russian Federation) that had received intensive training and supervision to promote warm, caring, sensitive, and responsive caregiver–child interactions versus a similar orphanage that did not.

For this project, 5 coders were trained by the scale's originators (R.B.M., C.J.G.). Training consisted of approximately 9 hr spread over 3 days of discussing each rating to come to a common definition, observing caregiving activities that had been taped in different orphanages and U.S. childcare centers, and one morning of "live" practice observation in a preschool center in which pairs of coders individually rated the behaviors and then compared and discussed their ratings.

Each of the 5 coders was then paired with each of the other coders to produce 10 pairs of coders, who collectively rated 14 different caregivers in the target orphanage as part of the pre-intervention assessments (One coder's data were randomly chosen to be used in the pre-intervention dataset.) Over all 18 items for each pair, coders were identical in the rating they assigned on 56% of the items and were within 1 point on 91% (Groark, McCall, & the Whole Child International Team et al., 2007). Percent agreement was slightly higher for the feeding activity than it was for bathing/dressing or free-play episodes, but percent agreement did not differ as a function of whether the caregiver served children birth to 3 years or 4 to 8 years of age. Coders were relatively uniform in their agreement with the other coders. The same 5 coders conducted both the pre- and postintervention CCSERRS assessments.

Battelle Developmental Inventory (BDI, Spanish version; Newborg, 2005). Children's general behavioral development was assessed with the BDI. Three assessors were trained to conduct the BDI, and then intercoder reliability was evaluated by having the 3 coders simultaneously assess 10 orphanage children between 2 and 64 months of age as part of the pre-intervention assessment (Again, data from one assessor were randomly selected for the pre-intervention dataset.) For each child, 1 coder administered the assessment, and all 3 coders, independently of the others, scored the child's performance. Each coder was the test administrator for 3 or 4 children.

Over these pairs, the 3 coders differed an average of 2.03 in total DQ and had a Kendall's Coefficient of Concordance of .99. Concordance values for the five subscales ranged between .75 (for Communication) and .95 to .99 for the other four subscales. The intraclass correlation for Total Score was .94, and those for the subscales were Adaptiveness (.83), Personal-Social (.96), Communication (.97), Motor (.67), and Cognition (.80). Further, an analysis of variance over all pre-intervention assessments, not just the 10 reliability cases, comparing caregivers, coders, and activity (ignoring the lack of independence associated with some caregivers being coded by more than 1 coder), showed that 38% of the variance was associated with differences between caregivers, 8% with activity, and only 1% with coders, indicating that different coders accounted for very little of the total variance relative to caregivers (Groark et al., 2007). Two of the original 3 coders who administered the pre-intervention BDI conducted the postintervention BDI assessments.

RESULTS

Data analyses focused on pre- versus postintervention differences with respect to the general environment of the wards, the socioemotional-relationship character of caretaker-child interactions, and children's general behavioral development.

Ward Environment (ITERS/ECERS)

Wards for both younger (<36 months) and older children scored higher at postintervention than they did at pre-intervention.

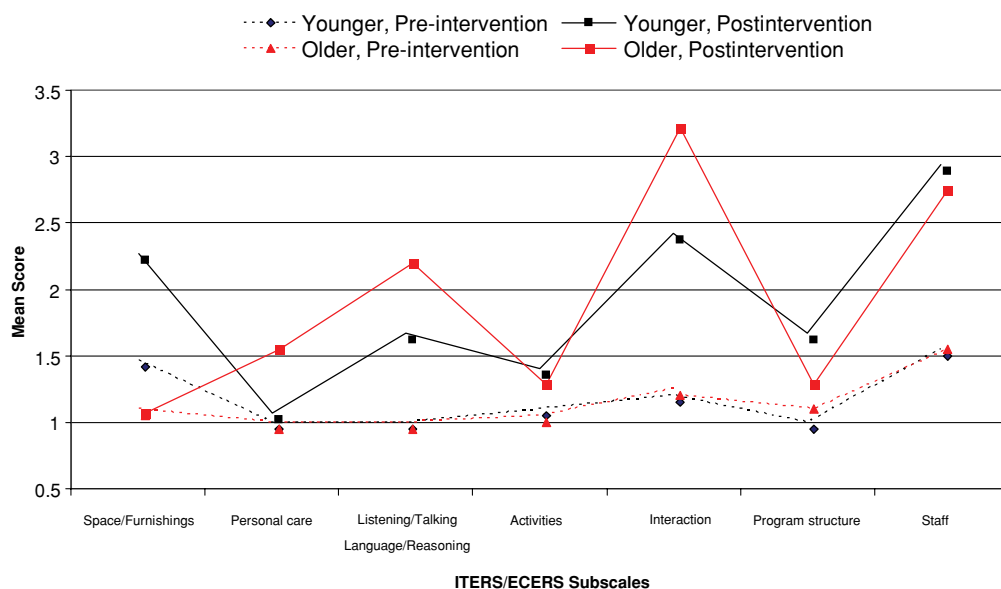


FIGURE 1. ITERS/ECERS scores at pre- versus postintervention for wards with younger versus older children.

Total Scores. The mean ITERS for the younger groups rose from 1.21 to 1.83, the mean for the older wards increased from 1.14 to 1.83, and all six groups individually improved over their pre-intervention assessments ($p < .02$).

Subscale Scores. Figure 1 presents ITERS/ECERS subscale scores for the younger and older wards at pre- and postintervention. The small number of wards precluded statistical analyses, but it is clear that the larger improvements were made on the Listening/Talking or Language/Reasoning subscale (mainly the individual items of Helping Children Understand Language, Using Language to Develop Reasoning Skills, and Informal Use of Language), the Interaction subscale (mainly the individual items of Discipline, Staff–Child Interactions, and Peer Interactions), and the Staff subscale (because all staff had received training). Minimum improvements were made on Space and Furnishings (except for the youngest groups because of the addition of the Pikler changing tables and bookcases to hold toys), Personal Care, Activities, and Program Structure.

Conclusion. These results are consistent with the intervention philosophy that focused on sensitive and responsive one-on-one caregiver–child interaction during routine caregiving activities in a consistent manner emphasizing caregiver–child engagement, child-directed interaction, much talking, and warm, caring, responsive interactions. However, additional, fine-grained analyses revealed these subscales and individual-item improvements were inconsistently manifested across different wards, suggesting that the intervention may have had greater influence on some caregivers than it did on others.

CCSERRS

The results for the CCSERRS are presented for two nonindependent samples.

TABLE 2. Caregiver–Child Social/Emotional/Relationship Rating Scale (CCSERRS) Total Scores at Pre- and Postintervention for Feeding, Changing/Bathing, and Free-Play Activities for the Quasi-Cross-Sectional and Longitudinal Samples

	Feeding			Changing/Bathing			Free Play		
	Pre	Post	Gain	Pre	Post	Gain	Pre	Post	Gain
Younger Children									
Quasi-Cross-Sectional	22.3 (8.0)	34.7 (8.4)	12.4	22.8 (6.3)	39.0 (6.4)	16.2	26.9 (8.3)	39.0 (6.4)	8.2
Longitudinal	20.3 (5.5)	37.4 (7.6)	17.1	24.4 (4.9)	38.9 (6.9)	14.5	22.1 (6.6)	38.9 (6.9)	11.8
Older Children									
Quasi-Cross-Sectional	18.4 (9.3)	25.9 (6.6)	7.5	20.1 (8.1)	24.8 (9.0)	4.7	24.0 (9.5)	24.8 (9.0)	5.4
Longitudinal	17.1 (9.1)	24.8 (5.6)	7.7	18.3 (7.8)	26.0 (7.1)	7.7	25.6 (6.2)	26.0 (7.1)	2.5

Quasi-Cross-Sectional $n_s = 12$ (Younger), 20 (Older); Longitudinal $n_s = 9$ (Younger), 11 (Older). Maximum score = 54. Posttest means are in boldface for contrast.

The *quasi-cross-sectional sample* included all caregivers working on these wards at pre-intervention and all caregivers working at postintervention, only some of whom were included in both pre- and postintervention assessments. Analyses treated pre- versus postintervention as independent groups despite this partial overlap, which is a statistically conservative strategy. This sample has the advantage of maximizing the number of cases available and giving a portrait of the entire orphanage before and after the intervention, but it has the limitation that pre-post changes may be associated with selective attrition or hiring of caregivers between pre- and postintervention.

The *true longitudinal sample* includes only caregivers who were present at both pre- and postintervention assessments. This sample has the advantage of displaying true intra-individual change in caregiver behaviors, but it has the limitations of a smaller n and potentially selective retention of caregivers. Presumably, if results are similar between samples—which they were—they converge on describing the effect of the intervention.

Quasi-cross-sectional sample. A Pre- versus Postintervention \times Younger versus Older Wards \times Activity (feeding, bathing/changing, free play) analysis of variance on CCSERRS Total Scores revealed a main effect for activity, $F(2, 197) = 3.34, p = .037, \eta^2 = .03$, and a Younger/Older \times Pre-/Postintervention interaction, $F(1, 197) = 7.74, p = .006, \eta^2 = .04$ (see Table 2). But the largest effect was for pre- versus postintervention, $F(1, 197) = 61.82, p = <.001, \eta^2 = .24$.¹

Thus, these results indicate that CCSERRS scores predominantly improved from pre- to postintervention, but not equally for each activity or age group.

Specifically, simple effects tests showed that caregivers in both the younger, $F(1, 197) = 58.10, p < .001, \eta^2 = .23$, and older wards, $F(1, 197) = 12.60, p < .001, \eta^2 = .06$, improved significantly from pre- to postintervention, but the partial effect size for the younger wards was nearly four times larger than that for the older wards. Further, simple effects tests for each

¹Note that η^2 is the partial eta-squared effect size for each source of variance with the other sources partialled out as provided by SPSS. Thus, it is not Cohen's effect size, and the total effect of the intervention, for example, might include both the main effect plus some parts of the interaction, depending on its form.

activity were conducted, and improvement was significant for both younger and older wards for feeding, $F(1, 77) = 24.09, p < .001, \eta^2 = .24$; $F(1, 77) = 7.50, p < .001, \eta^2 = .09$, respectively, although the effect size was nearly three times larger for the younger wards. Further, only the younger wards improved significantly for changing/bathing, $F(1, 58) = 35.32, p < .001, \eta^2 = .38$; $F(1, 58) = 2.82, p = .099, \eta^2 = .05$, respectively, and for free play, $F(1, 62) = 7.34, p < .009, \eta^2 = .11$; $F(1, 62) = 3.28, p = .075, \eta^2 = .05$, respectively.

True longitudinal sample. Although the number of cases for the true longitudinal sample is small (9 caregivers for younger children and 11 for older children), the results were similar, and effect sizes were larger for this sample than they were for the quasi-cross-sectional sample.

There was a significant effect for pre- to postintervention improvement, $F(1, 18) = 75.02, p < .01, \eta^2 = .81$, with a very large effect size (.81). In addition, there was a significant interaction of this improvement with age, $F(1, 18) = 12.90, p = .002, \eta^2 = .42$. Simple effects tests showed caregivers for both younger and older wards improved from pre- to postintervention, $F(1, 18) = 68.24, p < .001, \eta^2 = .79$; $F(1, 18) = 14.28, p = .001, \eta^2 = .44$, respectively, but the effect size was nearly twice as great for the caregivers of younger children than it was for older children.

Conclusion. As with the ITERS/ECERS, the results for the CCSERRS reflect the predominant emphases of the intervention program. The intervention stressed warm, caring, sensitive, and responsive caregiver–child interactions primarily during feeding and bathing/changing, and the training was especially relevant for caregivers of children birth to 3 years. Consequently, CCSERRS scores showed greater intervention-effect sizes for caregivers serving younger children than they did for older children, and caregivers for both age groups improved their interactions during feeding, but only caregivers for younger children improved for changing/bathing and free play. Intervention effect sizes were substantial for caregivers of younger children (e.g., .79 for the longitudinal sample).

BDI

The BDI scores were analyzed separately for a quasi-cross-sectional sample of all children available at pre-intervention and all children at postintervention plus a true longitudinal sample of all children who were assessed at both pre- and postintervention. In addition to the advantages and limitations of these two samples described earlier, the longitudinal sample consisted only of children who received the maximum possible exposure to the intervention in their respective wards. Separate analyses within these groups were performed on Total Scores and the set of five BDI subscales as a multivariate set.

Quasi-cross-sectional sample, Total Score. Results for Total Scores (DQs) for the quasi-cross-sectional sample are presented in Table 3, with Total DQ at the top. Both younger and older children improved from pre- to postintervention (+12.37, +14.62 DQ points, respectively), with a combined improvement of 13.51. The interaction was not significant, $F(1, 110) = .10$.

A Pre-versus Postintervention \times Younger versus Older Children analysis of variance produced main effects for pre- versus postintervention, $F(1, 110) = 40.67, p < .001, \eta^2 = .27$, as well as a smaller effect for age, $F(1, 110) = 13.13, p < .001, \eta^2 = .11$, which indicates that the older children scored more poorly at both pre- and postintervention than did the younger children. Simple effects tests confirmed that both younger, $F(1, 110) = 17.06, p < .001, \eta^2 = .13$,

TABLE 3. Means (SDs) for Battelle Developmental Inventory Developmental Quotients (DQ) for Younger and Older Children Pre-/Postintervention for the Quasi-Cross-Sectional Sample

BDI Scale	Pre-Intervention	Postintervention	Change
Total DQ			
Younger children (<i>n</i> = 26, 26)	67.95 (8.62)	80.32 (14.35)	+12.37***
Older children (<i>n</i> = 35, 27)	59.94 (7.38)	73.56 (12.34)	+14.62***
Adaptive DQ			
Younger children	75.52 (13.67)	87.71 (13.67)	+12.19**
Older children	67.49 (11.64)	72.84 (17.04)	+5.35
Personal–Social DQ			
Younger children	75.78 (10.14)	87.68 (12.89)	+11.90***
Older children	67.90 (8.38)	79.96 (13.61)	+12.06***
Motor DQ			
Younger children	82.02 (12.16)	86.21 (14.03)	+4.19
Older children	74.08 (10.75)	87.59 (15.32)	+13.51***
Communication DQ			
Younger children	61.89 (7.30)	73.15 (14.60)	+11.26***
Older children	62.99 (9.97)	74.13 (14.73)	+11.14***
Cognition DQ			
Younger children	73.61 (10.68)	84.95 (13.51)	+11.34***
Older children	61.29 (7.82)	72.00 (11.96)	+10.71***

Note. Significance tests are descriptive and approximate because a minority of children are represented in both pre- and postintervention groups.

p* < .01. *p* < .001.

and older children, $F(1, 110) = 24.27, p < .001, \eta^2 = .18$, improved pre- to postintervention, but effect sizes were modest.

Table 4 displays the percentage of children with Total DQs between certain values at pre- vs. postintervention. Although older children scored lower at each assessment, the relative trends

TABLE 4. Percentage of Children Scoring Between Certain Total Developmental Quotient (DQ) Values at Pre- Versus Postintervention for the Quasi-Cross-Sectional Sample

	DQs				
	<70	70–79	80–89	90–99	100+
Pre-Intervention	82.0%	13.1%	4.9%	0%	0%
Postintervention	27.8%***	27.8%†	29.6%***	11.1%*	3.7%

Note. Significance tests between pre- and postintervention within each DQ category are descriptive and approximate because a minority of children were in both assessments; comparisons between categories are not independent.

†*p* < .10. **p* < .05. ****p* < .001.

across score values were similar for the two age groups. Not surprising but important, the intervention had its most substantial effect on reducing the percentage of children who scored very poorly. Although 82.0% of children scored below 70 at pre-intervention (i.e., two *SDs* below the U.S. parent-reared norms for which only 2% would be expected to score this low), only 27.8% of the orphanage children did so after the intervention. In addition, no child scored above 90 at pre-intervention, but 14.8% did so after the intervention. Note that these data are for the quasi-cross-sectional sample, which means that some children who arrived later would not have been exposed to the entire 4+ months of the intervention.

Comparability of samples. For the longitudinal sample, children's pre-intervention Total DQs were compared with pre-intervention Total DQs for the quasi-cross-sectional sample. These means were nearly identical for the two samples for younger children (68.00 and 67.95, respectively) and older children (59.61 and 59.94, respectively), indicating that selective attrition probably did not influence which children remained in the longitudinal sample for the duration of the study.

Longitudinal sample, Total Score. Results for the true longitudinal sample were quite similar to those of the quasi-cross-sectional sample, and effect sizes were larger, as is typical of longitudinal analyses. In addition to the independent variable of pre- versus postintervention, there were three age groups composed of (a) children who stayed within a younger age group ($n = 6$) from pre- to postintervention, (b) those who made the transition from younger to older ($n = 9$) wards, and (c) those who remained in an older ward ($n = 17$). This analysis of variance predominantly produced a pre- versus postintervention effect, $F(1, 29) = 20.18, p < .001, \eta^2 = .41$, and a smaller effect for age, $F(2, 29) = 3.32, p = .05, \eta^2 = .19$. The interaction was not significant, $F(2, 29) = 1.61, p = .22$.

The interaction was not significant, but simple effects tests for the three age groups revealed an important difference. Children who remained in the younger wards improved from pre- to postintervention, 69.01 to 83.81 = +14.79; $F(1, 29) = 7.68, p = .01, \eta^2 = .21$, as did children who remained in the older wards, 59.61 to 73.77 = +14.16; $F(1, 29) = 19.96, p < .001, \eta^2 = .41$. Conversely, children who transitioned from a younger to an older ward did not improve significantly, 67.33 to 72.45 = +5.12; $F(1, 29) = 1.38, p = .25, \eta^2 = .05$. Specifically, all 6 of the younger children improved (3 between 1.4 and 1.6 *SDs*), 15 of 17 older children improved (5 more than 1.4 *SDs* and 1 as much as 2.8 *SDs*), but only 5 of 9 children who transitioned from younger to older wards improved (2 between 1.4 and 1.6 *SDs*, but 4 actually declined up to 1.3 *SDs*). Thus, the intervention produced nearly three times as much improvement in DQs and had from four to eight times the effect size on children who remained in the same wards during the intervention period than it did for children who changed wards, caregivers, and agetates.

Quasi-cross-sectional sample, subscales. Multivariate analyses of variance (MANOVA) analogous to those described earlier were conducted on the five BDI subscales. Again, results for the two samples were quite similar, and with the exception of one or two subscales, the results were similar to those described earlier for Total DQ.

For the quasi-cross-sectional sample, the MANOVA produced significant effects for pre- versus postintervention, $F(5, 106) = 9.55, p < .001, \eta^2 = .31$, and age (younger vs. older), $F(5, 106) = 15.54, p < .01, \eta^2 = .42$, but in contrast to the results for Total DQ, the interaction was significant, $F(5, 106) = 2.51, p = .034, \eta^2 = .11$, although no univariate interaction was

significant. The pre- versus postintervention change was significant for all five subscales, $F_s = 10.01\text{--}31.84$, $\eta^2 \eta_s^2 = .08\text{--}.22$, with the largest effect size for the Personal-Social subscale (.22), which was the main focus of the intervention, followed by the Cognition subscale (.20). Older children generally scored lower, and pre- versus postintervention improvements were significant only for the Adaptive, Personal-Social, and Cognition subscales. Subscales and multivariate simple-effects tests demonstrated that both younger and older children improved with similar effect sizes (.24 and .20, respectively).

Longitudinal sample, subscales. The results for the true longitudinal sample were very similar to the quasi-cross-sectional results and similar to the results for Total DQ. There was general improvement from pre- to postintervention, $F(5, 25) = 3.25$, $p = .02$, $\eta^2 = .39$, and the interaction with age was significant, $F(10, 50) = 2.27$, $p = .028$, $\eta^2 = .31$. This interaction reflected the fact that although children who remained in younger wards improved, $F(5, 25) = 2.44$, $p = .06$, $\eta^2 = .33$, and those who remained in older wards improved, $F(5, 25) = 5.12$, $p = .002$, $\eta^2 = .51$, children who made a transition from younger to older groups did not, $F(5, 25) = 1.23$, $p = .34$, $\eta^2 = .20$. Thus, children who remained in their age group had effect sizes that were 50 to 150% larger than those who made this transition. Univariate results were similar: Younger children improved significantly or nearly so ($p < .07$) on all subscales except the Communication subscale, older children improved on all subscales except the Adaptation subscale, and children who made an age-group transition did not improve significantly on any subscale.

Conclusion. Generally, children improved an average of 13.51 Total Score DQ points, younger and older children improved approximately equally, improvements occurred for nearly all subscales, and the intervention reduced the percentage of children scoring 70 or below from 82 to 27.8%. However, children who made a transition from a younger to an older ward, experiencing new caregivers and new and older peers, did not improve significantly.

DISCUSSION

Summary

The intervention. This intervention in a depressed orphanage in Latin America emphasized one-on-one warm, caring, sensitive, and responsive caregiver-child interactions, especially during routine-caregiving chores (e.g., feeding, bathing/changing), plus some increased consistency of caregivers. The intervention was targeted primarily at infants and toddlers in wards that were for children under 36 months of age, and caregivers in these wards received technical assistance in implementing the training in their routine-caregiving behaviors. In addition, unanticipated events produced a mixed intervention for older children, many of whom were denied access to preschool services during baseline because a child wandered away from the offsite preschool facility and attendance was temporarily stopped, followed later by a return to offsite and onsite preschool and recreational activities.

Ward and caregiver changes. Pre-post differences in the ward environments and caregiver-child interactions reflected the nature of the intervention. Wards improved on environmental rating scales, especially with respect to listening/talking or language/reasoning and caregiver-child interaction. Further, there were more warm, caring, sensitive, responsive, and engaging

caregiver–child interactions at postintervention. The pre–post differences were greater for caregivers serving younger children: Caregivers for younger children improved their interactions with children during all three activities (feeding, bathing/changing, free play), but caregivers for the older children had higher postintervention scores only during the feeding episodes.

Children's development. Despite different interventions for younger and older children, pre–post differences in general development were similar for younger and older children—an average of 13.5 DQ points at posttest. Although significance levels varied somewhat, such differences occurred over all subscales of the BDI. Older children started at lower developmental levels, perhaps because of a longer residency in an unstimulating orphanage plus the temporary absence of preschool activities during the pre-intervention assessment, but the pre–post differences were approximately the same number of DQ points as those for the younger children, who were the primary focus of the intervention.

Unexpected Results

Generally, the results were consistent with expectations, except for two aspects.

Older and younger children show same pre-post differences. Given that the training pertained primarily to children birth to 3 years and that wards serving children at these ages had fewer and more consistent caregivers who received technical assistance whereas wards serving older children did not, it is surprising that the pre–post differences in BDI-IQs of older children were as large as those for younger children. The simple interpretation is that training, which was similarly provided to caregivers serving children of all ages, was sufficient to produce the apparent developmental gains in all age groups. Further, the amount of technical assistance was rather minimal, perhaps too little, to produce better performance in the younger groups.

However, training alone (i.e., without technical assistance or supervision) has frequently been demonstrated to be relatively ineffective at producing caregiver behavioral change (Kelley, 1999; St. Petersburg–USA Orphanage Research Team, 2008). Further, environments and caregivers in this study improved to a greater extent in younger versus older wards. However, it may be that the intervention's emphasis on one-on-one caregiver–child interaction during routine-caregiving duties with a great deal of reciprocal conversation may actually be disproportionately beneficial to somewhat-older children who have a greater command of language than do younger children. Further, older children had much lower pre-intervention BDI DQs than did younger children, possibly reflecting a cumulative deficit as a result of living in the unstimulating orphanage for a longer period of time and because preschool services were discontinued before pre-intervention assessments were conducted. Preschool services as well as the opportunity for swimming, soccer, and arts activities and interactions with nonorphanage peers were introduced for older children before postintervention assessments and may have contributed to older children's posttest BDIs.

Collectively, then, it may be that for children who are severely delayed, training caregivers to be more sensitive, responsive, talkative, and engaging may have a positive benefit on the development of children birth to 8 years—a proposition consistent with the literature on orphanage interventions (Rosas & McCall, 2009).

Children who changed wards may have benefited less. Although the statistical evidence was somewhat inconsistent, children who made a transition from younger to older wards during the course of the study improved less than did children who remained in either younger or in older wards. Presumably, transitioning to a new room, new caregivers, and new and older peers was at least temporarily disruptive to developmental improvement.

This result is the first direct evidence suggesting that the common practice of “graduating” children periodically from one homogeneous age group to another may be detrimental to children’s development. This practice is often justified on practical grounds: It’s easier to take care of groups of children who are the same age, they have same-aged peers to play with, and they are at less risk of older and larger children dominating or physically hurting them. But in the history of mankind, young children were always reared in heterogeneous age groups (i.e., families) until institutional rearing emerged (Hartup, 1976; Konner, 1975). Attachment theorists would caution that graduations increase the number of different caregivers and the inconsistency of caregiving behaviors in children’s lives; the current data suggest that graduations may impede children’s general developmental progress; and the St. Petersburg project showed that graduations could be eliminated in an orphanage and that heterogeneous age groups could be introduced with no problems and many benefits (St. Petersburg–USA Orphanage Research Team, 2008).

The Intervention and Its Implementation

This study represented a pilot for Whole Child International to implement such an intervention, and as such, it is a case study of intervening in an entire institution. As is common in field studies, the implementation did not always proceed smoothly, in part because of unexpected and uncontrollable events and in part because of limited resources and planning.

For example, the directorship of the orphanage changed hands shortly before the intervention was to be implemented. Thus, the new director inherited this project, a circumstance that may contribute to less “buy in” and enthusiasm by the director—a factor often cited as being crucial to implementation success (St. Petersburg–USA Orphanage Research Team, 2008; Taneja et al., 2002).

Second, during implementation, it was clear that more technical assistance, monitoring, and encouragement of caregivers to incorporate their training in their ward behavior was needed, and this need was only partly fulfilled. For example, there was minimal local-staff supervision, and the two local specialists recruited to provide technical assistance may not have been sufficiently trained and were sometimes at odds with formal-staff supervisors. Further, no technical assistance was given to caregivers serving older children. Technical assistance in the form of continuous monitoring of caregiver behavior and encouraging the implementation of what was trained in day-to-day behavior on the ward with the children is crucial to behavioral change (Kelley, 1999; St. Petersburg–USA Orphanage Research Team, 2008).

Third, an influx of new children meant that group size could not be reduced as originally planned. For example, the child-to-caregiver ratio remained high—approximately 5 to 11:1 depending on how much help the aides provided with routine caregiving—a figure that is high by U.S. childcare standards for infants and toddlers.

Fourth, although the intervention produced rather intensive one-on-one caregiver–child interactions during feeding and bathing, there was relatively less caregiver attention or interaction with children during free-play times, a theme consistent with the Pikler Institute philosophy. Had groups been made heterogeneous by age, which the orphanage staff resisted, there might have

been more time for caregiver–child interaction during play time (St. Petersburg–USA Orphanage Research Team, 2008).

Fifth, in the future, several other factors important to the successful implementation of such an intervention (Fixsen et al., 2005; Groark & McCall, 2008; St. Petersburg–USA Orphanage Research Team, 2008) might be considered, including preparing administrators and staff for the intervention with a logic model approach, incorporating more local supervision into the intervention, implementing more structural changes that support caregiver–child relationship building (e.g., smaller groups and child-to-caregiver ratios), emphasizing sensitive and responsive caregiver–child interactions in contexts in addition to feeding and bathing/changing (e.g., play, group activities), and planning more age-appropriate activities for older children and their caregivers.

Nevertheless, across these irregularities, the results were quite clear and cohesive in showing that postintervention ward environments, the sensitivity and responsivity of caregiver–child interactions, and children’s developmental levels were higher than those at pretest. Further, although warm, sensitive, responsive caregiver–child interactions are often emphasized for children under 3 years, they benefited children up to 8 years by the same amount.

Potential confounds. But can these improvements be attributed to the intervention, and if so, which parts of the intervention? This study did not have random assignment or a no-treatment comparison group, so causal inferences cannot be made with confidence. No-treatment comparison groups in field studies, especially pilot studies, are expensive, difficult to obtain, and provide nearly no benefits to the no-treatment institution. Nevertheless, it seems unlikely that the observed changes occurred spontaneously because a new director took over or were a consequence of repeated testing of caregivers and children. Specifically, the pre-post differences in caregiver ward behavior corresponded to the areas emphasized by the training, not general improvement; the CCSERRS contains a check on caregivers “performing” for the observers; and CCSERRS results for two other untreated orphanages assessed later in another Latin American country were even lower than those of the pre-intervention scores reported here (McCall, Groark, Fish, & The Whole Child International Team, in press). In addition, children were posttested at least 1 year after pretesting, and the number of repeatedly tested longitudinal children was small relative to the cross-sectional sample. This means most posttest children were not pretested, which obviates a repeated testing benefit for most children. Further, in our judgment, the new director who inherited the project was less, not more, enthusiastic about the intervention than was the original director with whom the intervention was planned. Thus, it seems unlikely that the pre-post differences reported here are simply the result of secular change or repeated testing. Finally, the observed pre-post differences were largely consistent with the nature of the intervention and with the literature on stimulus interventions in orphanages (Rosas & McCall, 2009).

Field interventions of this sort are almost always clusters of elements, the effects of which cannot be teased apart. Initially, it is reasonable to improve as many aspects in as many ways as possible to demonstrate the potential for change; but ultimately, if interventions are to be brought to scale, they will need to be reduced to their essential components to be affordable and efficient.

Magnitude of Changes

Relative to the literature (Rosas & McCall, 2009), the current intervention was intermediate in intensity and produced an intermediate level of benefit for the children. Although environmental

and caregiver–child ratings improved and effect sizes were moderate-to-large, the extent of these improvements was modest.

ITERS/ECERS scores were 1.14 at pretest and only 1.83 at posttest on a scale of 1 to 7 (best). Although CCSERRS scores improved for caregivers serving young children, they did so only approximately .7 on a scale of 0 to 3 (best), and their average posttest score was approximately 2.0 of a maximum of 3.0, leaving room for further improvement. Similarly, the effect size was large, but the actual pre-posttest difference was 13.5 DQ points. Although 82% of children had DQs below 70 before the intervention and only 27.8% did so afterward, more than half of the children still had DQs less than 80. It may be that in highly depriving orphanages, even relatively modest improvements in caregiver behavior can produce noticeable improvements in children (Rosas & McCall, 2009).

At the same time, more substantial interventions have produced much greater benefits. In a very large, comprehensive, and intense quasi-experimental intervention following the same general philosophy of improving caregiver–child socioemotional interactions and reducing the number of different caregivers in orphanages in St. Petersburg, Russian Federation, very substantial improvements in caregiver behavior on the ward were obtained (McCall et al., 2010; St. Petersburg–USA Orphanage Research Team, 2008), with correspondingly substantial improvements in child DQs (e.g., from a mean DQ of 57–92, although the standardization sample and scoring scale were different than those used here).

Conclusion

This study is at least consistent with the very few others that have attempted to permanently change an entire institution and its regular staff, especially to increase warm, sensitive, responsive caregiver–child interactions (Rosas & McCall, 2009; St. Petersburg–USA Orphanage Research Team, 2008). These studies are consistent with the propositions that an entire institution can be changed, children’s development can improve, and the focus on socioemotional caregiver–child interactions alone can produce gains in all aspects of children’s development.

These studies converge on the hypothesis that the relative lack of warm, caring, sensitive, and responsive caregiver–child interactions in an orphanage context is an exceedingly corrosive early experiential deficit. Furthermore, without changing nutrition, medical care, and child–caregiver ratios (which in this study remained at a high level), such a socioemotional intervention can be associated with meaningful improvement in children’s development, and the more comprehensive and intensive the improvement in these caregiver–child interactions, the better the development of the children in these institutions.

REFERENCES

- Brossard, M., & Decarie, T.G. (1971). The effects of three kinds of perceptual–social stimulation on the development of institutionalized infants: Preliminary report of a longitudinal study. *Early Child Development and Care*, 1, 111–130.
- Casler, L. (1965). The effects of extra tactile stimulation on a group of institutionalized infants. *Genetic Psychology Monographs*, 71, 137–175.
- Fixsen, D.L., Naoom, S.F., Blasé, K.A., Friedman, R.M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa: University of South Florida, Louis de la Parte Florida Mental Health Institute, National Implementation Research Network (FMHI Publication No. 231).

- Gavrin, J.B., & Sacks, L.S. (1963). Growth potential of preschool-aged children in institutional care: A positive approach to a negative condition. *American Journal of Orthopsychiatry*, 33, 399–418.
- Gonzalez-Mena, J. (2005). *Una guía para las rutinas cotidianas del cuidado infantil, segunda edición* [A guide to daily routines of child care, 2nd ed.]. Sacramento: California Department of Education and WestEd.
- Groark, C.J., & McCall, R.B. (2008). Community-based interventions and services. In M. Rutter, D.V.M. Bishop, D.S. Pine, S. Scott, J. Stevenson, E. Taylor, & A. Thapar (Eds.), *Rutter's child and adolescent psychiatry* (5th ed., pp. 971–978). London: Blackwell.
- Groark, C.J., McCall, R.B., & the Whole Child International Team. (2007). *Latin American orphanage baseline report*. Pittsburgh, PA: University of Pittsburgh, Office of Child Development.
- Hakimi-Manesh, Y., Mojdehi, H., & Tashakkori, A. (1984). Short communication: Effects of environmental enrichment on the mental and psychomotor development of orphanage children. *Journal of Child Psychology and Psychiatry*, 25(4), 643–650.
- Harms, T., Clifford, R.M., & Cryer, D. (2005). *Early Childhood Environmental Rating Scale* (Rev. ed.). New York: Teachers College Press, Columbia University.
- Harms, T., Cryer, D., & Clifford, R.M. (2006). *Infant/Toddler Environmental Rating Scale* (Rev. ed.). New York: Teachers College Press, Columbia University.
- Hartup, W.W. (1976). Cross-age versus same-age peer interaction: Ethological and cross-cultural perspectives. In V. Allen (Ed.), *Children as tutors: Theory and research on tutoring* (pp. 41–55). New York: Academic Press.
- Kelley, S. (1999, September). Briefing paper: Regulatable factors in early childhood services. Pittsburgh, PA: University of Pittsburgh, Office of Child Development.
- Kim, P.I., Shin, Y.H., & White-Traut, R.C. (2003). Multisensory intervention improves physical growth and illness rates in Korean orphaned newborn infants. *Research in Nursing & Health*, 26, 424–433.
- Konner, M. (1975). Relations among infants and juveniles in comparative perspective. In M. Lewis & L. Rosenblum (Eds.), *Friendship and peer relations* (pp. 99–129). New York: Wiley.
- Lally, J.R. (Ed.). (1990). *Infant/toddler caregiving: A guide to social–emotional growth and socialization*. Sacramento: California Department of Education.
- Lally, J.R., Mangione, P.L., & Young-Holt, C.L. (Eds.). (1992). *Infant/toddler caregiving: A guide to language development and communication*. Sacramento: California Department of Education.
- McCall, R.B., Groark, C.J., & Fish, L.E. (2010). A Caregiver–Child Socioemotional and Relationship Rating Scale (CCSERRS). *Infant Mental Health Journal*, 31, 201–219.
- McCall, R.B., Groark, C.J., Fish, L.E., & The Whole Child International Team (2008). *Latin American orphanage intervention outcome report*. Pittsburgh, PA: University of Pittsburgh, Office of Child Development.
- McCall, R.B., Groark, C.J., Fish, L.A., & The Whole Child International Team. (in press). Characteristics of environments, caregivers, and children in three Latin American orphanages. *Infant Mental Health Journal*.
- Nelson, C.A. (2007). A neurobiological perspective on early human development. *Child Development Perspectives*, 1, 13–18.
- Nelson, C.A., Zeanah, C.H., Fox, N.A., Marshall, P.J., Smyke, A.T., & Guthrie, D. (2007). Cognitive recovery in socially deprived young children: The Bucharest early intervention project. *Science*, 318, 1937–1940.
- Newborg, J. (2005). *Battelle Developmental Inventory* (2nd ed.). Itasca, IL: Riverside.

- Rheingold, H.L. (1956). The modification of social responsiveness in institutional babies. *Monographs of the Society for Research in Child Development*, 21, Serial No. 63(2).
- Rosas, J., & McCall, R.B. (2009). *Characteristics of institutions, interventions, and children's development*. Pittsburgh, PA: University of Pittsburgh, Office of Child Development.
- Saltz, R. (1973). Effects of part-time "mothering" on IQ and SQ of young institutionalized children. *Child Development*, 44, 166–170.
- Sayegh, Y., & Dennis, W. (1965). The effect of supplementary experiences upon the behavioral development of infants in institutions. *Child Development*, 36, 81–90.
- Smyke, A.T., Dumitrescu, A.B., & Zeanah, C.H. (2002). Attachment disturbances in young children. I: The continuum of caretaking casualty. *Journal of the American Academy of Child & Adolescent Psychiatry*, 41, 972–982.
- Sparling, J., Dragomir, C., Ramey, S.L., & Florescu, L. (2005). An educational intervention improves developmental progress of young children in a Romanian orphanage. *Infant Mental Health Journal*, 26(2), 127–142.
- St. Petersburg–USA Orphanage Research Team. (2005). Characteristics of children, caregivers, and orphanages for young children in St. Petersburg, Russian Federation. *Journal of Applied Developmental Psychology: Special Issue on Child Abandonment*, 26, 477–506.
- St. Petersburg–USA Orphanage Research Team. (2008). The effects of early social–emotional and relationship experience on the development of young orphanage children. *Monographs of the Society for Research in Child Development*, 73, Serial No. 291(3).
- Taneja, V., Sriram, S., Beri, R.S., Sreenivas, V., Aggarwal, R., Kaur, R., & Puliyl, J.M. (2002). "Not by bread alone:" Impact of a structured ninety-minute play session on development of children in an orphanage. *Child: Care, Health and Development*, 28(1), 95–100.
- Tardos, A. (2007). *Bringing up and providing care for infants and toddlers in an institution*. Budapest, Hungary: Pikler-Loczy.
- UNAIDS, UNICEF, USAID. (2004). *Children on the brink: 2004. A joint report of new orphan estimates and a framework for action*. Washington, DC: USAID.
- Wolff, P.H., Dawit, Y., & Zere, B. (1995). The Solomuna Orphanage: A historical survey. *Social Science and Medicine*, 40, 1133–1139.
- Wolff, P.H., & Fesseha, G. (1998). The orphans of Eritrea: Are orphanages part of the problem or part of the solution? *American Journal of Psychiatry*, 155(10), 1319–1324.
- Wolff, P.H., & Fesseha, G. (1999). The orphans of Eritrea: A five-year follow-up study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 40(8), 1231–1237.
- Wolff, P.H., Tesfai, B., Egasso, H., & Aradom, T. (1995). The orphans of Eritrea: A comparison study. *Journal of Child Psychology and Psychiatry*, 36(4), 633–644.