

Quality of early child care and education: analyses at the ITERS-R indicator level

Barros S., & Peixoto C.

*School of Education, Polytechnic Institute of Porto, Porto, Portugal
Center of Psychology, University of Porto, Porto, Portugal
School of Education, Polytechnic Institute of Viana do Castelo, Viana do Castelo, Portugal*

Summary

The Infant/Toddler Environment Rating Scale - Revised Edition (ITERS-R; Harms, Cryer, & Clifford, 2005) is being used in several countries to measure quality of classrooms for children under 50 months of age. The main goal of the present research was to study the ITERS-R alternative scoring system. With this purpose, indicators were rated beyond the quality level assigned to each item in 160 classrooms, for children between 1 and 5 years of age, in the district of Porto. Results showed that this alternative scoring system allows researchers to describe classrooms' quality with more detail and to focus also on the classrooms' strong points. These results can also be useful to organize intervention in order to improve quality and, thus, have a positive effect in child development.

Introduction

The growing need of out-of-home early child care and education has been requiring more attention to such services' quality. Research has been showing that quality of care is related to children developmental outcomes (e.g., Bryant, Burchinal, Lau, & Sparling, 1994; National Institute of Child Health and Human Development Early Child Care Research Network, 2006). However, the assessment of child care quality and the definition of high-quality services have been the focus of many debates between researchers, parents and teachers.

The Infant/Toddler Environment Rating Scale - Revised Edition (ITERS-R; Harms, Cryer, & Clifford, 2005) has become a widely used instrument in research on the quality of programs for infants and toddlers in several countries. ITERS-R is based in research, professional values, and empirical knowledge, and is consistent with the Criteria for Quality Early Childhood Programs, defined by the National Association for the Education of Young Children (1998) and with the requirements of the Child Development Associate - CDA (CDA National Credentialing Program 1984 as cited in Harms, Cryer, & Clifford, 1999). Research has demonstrated that the ITERS-R correlates with other indicators of child care quality (e.g., Cost, Quality and Child Outcomes Study, 1995; Goelman et al., 2006).

In Portugal, assessments of quality with this instrument showed the existence of poor average quality and absence of good quality classrooms, as well as low variability of scores (Barros & Aguiar, 2010). Those analyses were performed based on data obtained in accordance to the scoring system recommended by the authors and generally used in research. Although previous results may contribute to substantiate the use of ITERS-R in the assessment of Portuguese toddler child care classrooms, mean scores at the item level concentrate at the lower end of the scale, hiding classrooms' adequate practices included in the higher indicators. These results mirror problems observed in child care centres but can also be related to item design and/or to strict requirements in the scoring system (Barros & Aguiar, 2010).

The main goal of the present research was to study the ITERS-R alternative scoring system.

Method

1. Participants

This study took place in the district of Porto in the north of Portugal. A stratified random sample of 80 non-profit private centers and 80 for-profit private centres was selected. In each of the 160 centers, we observed one classroom for children between 1 and 2 years old, between 2 and 5, or between 1 and 5. The average of children enrolled was 12.94 ($SD = 4.62$), with an average of 2.02 adults ($SD = 0.76$), and an average of 6.97 children per adult ($SD = 2.82$). In 82.5% of the classrooms, the lead adult (i.e., the adult responsible for providing or most of the direct work with children) was a trained teacher (i.e., a teacher with a college degree in early childhood education), and in 17.5% classrooms the lead adult was an untrained teacher (a teacher's assistant without a college degree). Untrained teachers reported between 5 and 16 years of

formal schooling and trained teachers had a bachelors' degree ($n = 52$), a licentiate degree ($n = 97$), or a post-graduate degree ($n = 5$). Adults' age ranged from 21 to 52 years ($M = 52.27$, $SD = 6.95$) and experience from 2 months to 52 years ($M = 8.09$, $SD = 6.66$).

2. Measures

Global classroom quality was assessed using the Portuguese translation of the *Infant/Toddler Environment Rating Scale - Revised Edition* (ITERS-R; Harms et al., 2005). The translation of ITERS-R used in this study was conducted by the first author, under the supervision of a senior researcher.

ITERS-R has 59 items, grouped into seven subscales: Space and Furnishings, Personal Care Routines, Listening and Talking, Activities, Interaction, Program Structure, and Parents and Staff. Scores on the ITERS-R range from 1 to 7 with 1 indicating *inadequate* quality, 5 indicating *minimal* quality, 5 indicating *good* quality, and 7 indicating *excellent* quality. Each item is presented as a 7-point scale, with descriptors/indicators for 1, 5, 5, and 7. In the traditional scoring system, the rating process is initiated by reading the indicators of 1 (*inadequate*) and moving to the following indicators if none of the inadequate indicators are present. If all of the indicators under an anchor are marked as present, the indicators under the next anchor are also scored. This method continues until an indicator is scored as not present. In that case, the indicators under that anchor are scored, but the observer stops scoring the item at that anchor, and the higher indicators are not scored. This is called the stop-rule by Hofer (2008). The total score is an average of the item scores.

In the alternative scoring system, all the 467 indicators are scored. In this study, for each of the items the percentage of indicators scored as present under the anchors of 5, 5 and 7 was calculated. The total score consists on the average of the percentage of indicators scored as present in each of the items.

Due to the presence of missing values, permitted by the scale, items 21, 25, 51, 52, and 56 were excluded from analyses. Cronbach's alpha coefficient on the overall scale was .85 for the traditional scoring system and .88 for the alternative.

Procedure

Training. Three observers were trained using the procedures and materials recommended by the ITERS-R authors. Training sessions were also conducted in 16 toddler child care classrooms; mean percentage of within-one interobserver agreement was 88%.

Data collection. Data were collected between September 2004 and October 2005. Each observer remained with the group of children

for at least 3 hours. Following the observation of each classroom, observers conducted a small interview with the lead teacher in order to score all indicators and to collect information on classroom structural features.

During the data collection procedure, interobserver agreement checks were conducted across 44 sessions (27.5% of the classrooms). Mean percentage of within-one interobserver agreement was 95.9 for the traditional scoring system and 88.68 for the alternative system; weighted k was 0.69 for the traditional scoring system and 0.61 for the alternative system.

Results

Using the traditional scoring system, overall mean results on the ITERS-R ranged from 1.62 to 4.09 ($M = 2.84$, $SD = 0.48$). Sixty one percent of classrooms were given scores that suggest inadequate quality (i.e., with an overall score under 3.00) and only 39% of classrooms were given scores that indicated minimal quality (i.e., with an overall score between 3 and 5). There were no high-quality classrooms.

Using the alternative scoring system, overall mean results on the ITERS-R ranged from 32.22% to 70.36% ($M = 53.20$, $SD = 8.24$). In a total of 159 classrooms, 60% had at least half of the indicators scored *Yes* (i.e., present) and almost 10% of classrooms had at least two thirds of the indicators scored *Yes*.

A strong association was found between mean results obtained through the traditional system and through the alternative system ($r = .94$, $p < .001$).

Following the traditional scoring system, item mean results ranged from 1.06 to 4.41. About half of the items (i.e., 18) presented mean scores that indicate the presence of minimal quality, while the remaining half presented mean scores that suggest the provision of poor-quality child care (see Table 1). From those 18 items with mean results lower than 3.00, 5 items scored higher than 50% with the alternative scoring system. This means that, although those 5 items had a score of 1 or 2, more than half of the indicators under the anchors of 3, 5 or 7 were present. However, 15 of the 18 items scored half than 50%, indicating that classrooms are failing in many of the requirements (see Table 1).

With the traditional scoring system, 21 items mean results were between 3.00 and 5.00 (i.e., minimal quality). Nineteen of these scored higher than 50% with the alternative scoring system (3 higher than 66%) and 2 scored lower than 50%. From these 21 items, all of the 7 items with average scores between 4.00 and 5.00 scored higher than 50% with the alternative scoring system.

	Traditional scoring stem		Alternative scoring	
	<i>M</i>	<i>SD</i>	♦ (%)	SL (%)
I. Space and Furnishings				
1. Indoor space	3.44	1.34	81.38	11.46
2. Furniture for routine care and play	3.16	1.61	66.67	15.73
3. Provision for relaxation and comfort	3.10	1.09	49.33	22.60
4. Room arrangement	3.16	1.31	65.14	18.12
5. Display for children	3.78	0.97	51.60	20.90
II. Personal Care Routines				
6. Greeting/departing	2.37	1.99	66.44	12.20
7. Meals/snacks	1.64	1.12	51.38	14.30
8. Nap	1.19	0.73	60.00	19.67
9. Diapering/toileting	1.06	0.31	45.44	12.28
10. Health practices	1.54	0.61	39.38	9.83
11. Safety practices	2.25	1.16	43.21	19.22
m. Listening and Talking				
12. Helping children understand language	4.24	1.33	70.00	16.37
13. Helping children use language	4.22	1.43	60.55	23.52
14. Using books	1.61	1.02	32.94	22.02
IV. Activities				
15. Fine motor	4.05	1.21	60.45	17.79
16. Active physical play	1.79	0.88	42.39	19.15
17. Art	3.11	1.36	69.90	20.09
18. Music and movement	2.39	1.24	55.69	14.56
19. Blocks	1.79	1.23	20.23	22.16
20. Dramatic play	3.51	1.41	55.09	25.17
21. Sand and water play	1.11	0.34	4.19	11.86
22. Nature/science	1.61	1.05	33.05	30.14
23. Use of TV, video and/or computer	1.48	0.77	25.13	14.22
24. Promoting acceptance of diversity	1.31	0.51	19.73	9.19
V. Interaction				
25. Supervision of play and learning	3.38	1.71	55.00	25.62
26. Peer interaction	4.36	1.06	57.92	15.52
27. Staff-child interaction	3.94	1.61	64.24	22.14
28. Discipline	3.14	1.32	53.06	19.82
VI. Program Structure				
29. Schedule	3.21	0.77	44.73	17.25
30. Free play	2.08	0.77	46.09	17.30
31. Group play activities	2.39	1.54	52.48	20.80
32. Provisions for children with disabilities	3.13	2.19	64.38	28.28
VII. Parents and staff				
33. Provisions for parents	3.33	1.18	60.34	13.77
34. Provisions for personal needs of staff	3.72	0.90	65.52	14.85
35. Provisions for professional needs of staff	4.26	1.65	73.59	18.17
36. Staff interaction and cooperation	4.16	1.17	68.63	12.39
37. Staff continuity	4.41	1.77	76.48	16.29
38. Supervision and evaluation of staff	2.41	1.28	27.92	24.16
39. Opportunities for professional growth	2.04	1.14	37.71	22.91

Table 1 Descriptives for the Items with the Traditional and the Alternative Scoring Systems

Conclusions

Results obtained with the traditional scoring system suggest that toddlers attending child care classrooms in the district of Porto are receiving poor quality care and education that does not appropriately respond to children's basic health, safety, and developmental needs (cf., Barros & Aguiar, 2010). The traditional scoring system assumes there are basic indicators that are necessary for any educational environment serving young children. If those indicators are not present, the score must correspond to inadequate quality, independently of other indicators that might be present beyond that low anchor.

The need to describe and understand what indicators child care centers are in fact meeting beyond the stop-rule (Hofer, 2008), and, at the same time, the need to study the alternative scoring system motivated this study.

The strong association between results obtained through traditional and alternative systems showed that, for research purposes, it might be sufficient to use the traditional one. Despite this strong correlation, the items where classrooms show the lowest and highest results are not exactly the same for both systems. For instance, 5 of the 18 items rated as inadequate quality had more than 50% of the indicators of the anchors 5, 5 and 7 scored Yes, but 2 of the 21 items that were scored as minimal quality had less than 50% of the indicators scored Yes. Using both systems can highlight areas where the classrooms, independently of the system we use, show the lowest results (such as Health practices, Using books, and many of the items on the Activities subscale). It is also possible to highlight items where classrooms show the best results (such as Helping children understand language, Staff-child interaction, and some items on the Parents and staff subscale).

The alternative scoring system allows researchers and practitioners: to obtain reliable data; to describe classrooms' quality with more detail and to identify the specific criteria at which the classrooms in general show more weaknesses and more strengths; to organize intervention in order to improve quality; and to focus also on the classrooms' strong points.

References

- BARROS, S., & AGUIAR, C., Assessing the quality of Portuguese child care programs for toddlers, *Early Child Res Q* 25 (4), 527-555, 2010.
- BRYANT, D. M., BURCHINAL, M., LAU, L. B., & SPARLING, J. J., Family and classroom correlates of Head Start children's developmental outcomes, *Early Child Res Q* 9 289-509, 1994.
- COST, QUALITY & CHILD OUTCOMES STUDY TEAM, *Cost, Quality, and Child Outcomes in child care centers, Public Report (2.ª edi iiQ)*, Denver,

- Economics Department, University of Colorado at Denver, 1995.
- GOELMAN, H., BARRY, F., KERSHAW, P., DOHERTY, G., LERO, D., & LANGRANGE, A., Towards a predictive model of quality in Canadian child care centers, *Early Child Res Q* 21, 280-295, 2006.
- HARMS, T., CRYER, D., & CLIFFORD, R. M., *In/ant/Toddler Environment Rating Scale - Revised Edition*, New York, Teachers College Press, 2003.
- HARMS, T., CRYER, D., & CLIFFORD, R. M. *In/anti Toddler Environment Rating Scale*, Retrieved July 2002, from www.fpg.unc.edu/~ecers/iters.html, 1999.
- HOFER, K. G., *Measuring quality in pre-kindergarten classrooms: assessing the Early Childhood Environment Rating Scale*, Unpublished doctoral dissertation, Vanderbilt University, 2008.
- NATIONAL INSTITUTE OF CHILD HEALTH AND HUMAN DEVELOPMENT EARLY CHILD CARE RESEARCH NETWORK, Child-care effect sizes for the NICHD Study of Early Care and Youth Development, *Am Psycho*, 61(2), 99-116, 2006.