Home Literacy Environment Profiles and Emergent Literacy of Preschool Children: A Cluster Analysis

Marina A. Kotzampasaki

Leiden University

Author Note

Marina A. Kotzampasaki, Social and Behavioral Sciences, Leiden University.

I am grateful to Katerina Kokkinou, Stamatina Marougkliani and Christina P. Theodosiou for the co-operation in the data collection and Zsofia K. Takacs for developing the HLE questionnaire and for personal assistance in manuscript preparation.

Abstract

This study investigated home literacy environment in relation to emergent literacy, through a cluster analytic approach. Fifty eight parents of kindergartens completed a HLE questionnaire, an author recognition test and a child book recognition test. Included were scores on literacy activities, watching television, emergent literacy as well as socioeconomic status. From the findings emerge three home literacy profiles differing in engagement in literacy activities: (1) high engagement in literacy activities and low television watching, (2) low engagement in any activity, and (3) low engagement in literacy activities but high in watching television. Socioeconomic status did not differ among clusters.

Keywords: emergent literacy, HLE, literacy activities, watching television, SES

Home Literacy Environment Profiles and Emergent Literacy of Preschool Children: A Cluster Analysis

Literacy development begins at early age, far before the start of formal instruction and it is defined as "Emergent literacy" (Whitehurst & Lonigan, 1998 ; Sulzby & Teale, 1996). That development depends on the literacy experiences of the child at home (Bus & van IJzendoorn, 1995; Davidse, de Jong, Bus, Huijbregts, & Swaab, 2010; Farver, Xu, Eppe, & Lonigan, 2006; Mol and Bus, 2011; Sénéchal, LeFevre, Hudson, & Lawson, 1996; Sénéchal, Pagan, & Lever, 2008). However, not all children have similar chances to develop literacy skills (Van Steensel, 2006; Chaney, 1994; Philips & Lonigan, 2009). Main aim of this study was to compile profiles of how children's literacy is supported in the home environment and how this varies with socio-economic status.

Emergent Literacy

It is assumed that literacy starts from an early age. "Emergent literacy" is a term that is used to describe the idea that literacy acquisition is a continuous development, rather than a skill that starts when children start school (Whitehurst & Lonigan, 1998). Emergent literacy is defined as the earliest phases of literacy development, from birth until children learn to read and write conventionally. Skills that are acquired before children become conventional readers and writers are named emergent literacy skills; they involve what is necessary for the development of reading and writing (Sulzby & Teale, 1996).

The "reading readiness" approach preceded emergent literacy and focused on identifying which skills are important to be mastered, so that children can best profit from formal reading instruction (Whitehurst &Lonigan, 1998). However, this perspective implies strict boundaries between formal reading and pre-reading period which cannot be made. The growing awareness that children familiarize with components of reading from an early age resulted in introduction of the term emergent literacy by Teale and Sulzby.

Whitehurst and Lonigan (1998) had conceptualized two domains for emergent literacy: *outside–in* and *inside-out* progress. Outside-in components represent children's comprehension of the concept that they are trying to read when inside-out components represent children's knowledge of rules for translating the writing into sound or sound into writing. Outside-in progress consists of four components: semantic, syntactic and conceptual language knowledge, the understanding and production of narratives, conventions of print, and emergent reading which is pretending to read and involves knowledge about the written register. Inside-out progress consists of five components: letter-name knowledge, phonological awareness, manipulations of syllables and manipulation of individual phonemes, repairing grammatical errors, letter-sound knowledge, and phonetic spelling.

Home Literacy Environment (HLE)

Home literacy environment (HLE) covers experiences that may contribute to *outside–in* and *inside-out* progress. Parents' engagement in literacy activities is often measured by self-report data on questions about the frequency of reading to the child, number of books available at home, the frequency of library visits, and playing word games with the child (Davidse, de Jong, Bus, Huijbregts, & Swaab, 2010; Farver, Xu, Eppe, & Lonigan, 2006; Sénéchal, LeFevre, Hudson, & Lawson, 1996; Sénéchal, Pagan, & Lever, 2008).

Especially, the frequency of shared-reading is a main component of the HLE. Print exposure is a main aspect of a literate environment (Bus & van IJzendoorn, 1995). Davidse, de Jong, Bus, Huijbregts, and Swaab (2010) found that children's knowledge of storybooks mediated the relation between HLE and literacy skills. A meta-analysis by Bus, van IJzendoorn and Pellegrini (1995) highlighted the importance of shared reading for language growth. Based on recent studies, the meta-analysis by Mol and Bus (2011) confirmed that print exposure predicted oral language skills as well as literacy skills.

Emergent literacy and HLE

Emergent literacy is an aspect strongly correlated with later reading achievement. Higher emergent literacy skills imply higher reading achievement in the early grades of primary education (Lonigan, Burgess and Anthony, 2000). Alphabetic knowledge such as identifying, labelling, and printing letters of the alphabet and printing and writing the proper name, are related to later reading achievement (Evans & Shaw, 2008).

The relation between emergent literacy and HLE has been the focus of numerous studies. Children's emergent literacy skills and language are predicted by responsiveness and support of the home environment (Roberts, Jurgens, and Burchinal, 2005). Emergent literacy is also correlated with family income and mothers' education (Whitehurst & Lonigan, 1998). Moreover, another aspect of HLE such as parents' literacy habits - parents' exposure to literacy -is related to book exposure of the children and to children's vocabulary (Sénéchal et al. 1996). Children's success in phonological tasks - another aspect of emergent literacy - is also related to family literacy (Chaney, 1994). Parents' teaching appears to be a significant contributor as well, to children's as print concepts, alphabet knowledge, invented spelling, and decoding (Sénéchal et al. 1996). Children's success in phonological tasks is related to family literacy (Chaney, 1994). Finally, print concepts including letter knowledge, orthographic awareness and word writing, are more improved by joint writing than joint reading (Aram & Biron, 2004).

In addition, numerous studies and meta-analyses focused more on shared reading as an aspect of HLE. Book exposure was found to be a significant contributor to emergent literacy, regardless of parental literacy. Moreover, numerous studies have shown the contributions of HLE to children's language development. The importance of children's book exposure for the development of vocabulary and listening comprehension has been shown (Evans & Shaw, 2008; Sénéchal & LeFevre, 2002). In addition, HLE was found to predict

vocabulary and letter knowledge outcomes (Davidse et al., 2010). Similar results of the positive correlation of shared reading with vocabulary morphological and syntax comprehension were revealed in other studies as well (Sénéchal et al., 2008). Moreover, book exposure throughout the preschool years may be an important influence on the development of children's vocabulary, listening comprehension, and phonological awareness (Sénéchal et al., 1996).

Television watching

Television viewing can be considered as an activity that may conflict with literacy-related activities. However, television watching might be a good replacement for literacy activities and contribute to language development. Studies that focused on the hypothesis that watching television contributes to language development revealed controversial results (Patterson, 2002; Philips & Lonigan, 2009; Scarborough & Dobrich, 1994). Patterson (2002) targeting bilingual toddlers found that television viewing was not related to vocabulary measures for both English- and Spanish-speaking children. However, it was significantly positively correlated with reading frequency for English but not for Spanish. Overall, Patterson (2002) concluded that watching television did not contribute to language growth, and in so far it did relations with reading frequency are inconsistent. Philips and Lonigan (2009) found that television watching was negatively related with reading frequency; engagement in watching television detracted the child from reading opportunities. Scarborough and Dobrich (1994) on the other hand, make a distinction between general television watching, including cartoons and non educational programs, and watching educational television programs such as *Sesame Street*. While general television watching was not related to any language growth, educational television was a contributor to language development and emergent literacy.

HLE profiles

Phillips and Lonigan (2009) identified variations in HLE through a cluster analytic approach. Their study resulted in 3 different clusters which were defined by three aspects: watching television frequency, engagement in direct literacy activities and engagement in indirect literacy activities. In line with the outside-in and inside-out model, they define direct literacy activities as the activities that focus on alphabet knowledge, phonological awareness and phonemic decoding capacities, while indirect literacy activities focus on language skills, vocabulary, understanding narratives.

The first cluster found was characterized by a high engagement in direct literacy activities and low engagement in indirect literacy activities and high frequency of television watching. The second cluster had low engagement in all aspects: watching television and direct and indirect literacy activities. The third cluster is characterized by low frequency of watching television and high engagement in both direct and indirect literacy activities.

So there is evidence that television watching is in conflict with literacy-related activities. However, the study did not support the distinction between direct and indirect activities as was reported by Sénéchal, LeFevre, Thomas, and Daley (1998); parents do not necessarily spend time on both types of activities but might mainly focus on book sharing or teaching literacy-related skills.

Socioeconomic Status (SES) and HLE

Numerous studies have shown that HLE is related to socioeconomic status (SES). Children from high SES families were found to have the most stimulating HLE (Van Steensel, 2006). Literacy activities are related to SES and more specific to family income and maternal education (Chaney, 1994). Moreover, literature for children varied among the SES

groups; in low SES families children own fewer children's books compared with high SES families (Chaney, 1994).

The above discussed study by Phillips and Lonigan, evidences that lower SES parents engage more in direct literacy activities while higher SES parents may or may not engage in direct and indirect activities (Philips & Lonigan, 2009). Moreover, the study demonstrates that children from lower SES families watch more television than children from higher SES families.

The Aim of This Study

The main aim of this study is to replicate the Phillips and Lonigan's (2009) findings, namely making clusters based on watching television and frequency of literacy activities. The clusters also include indicators for emergent literacy. The cluster analysis will be made based on 3 aspects: watching television, literacy activities and emergent literacy. In addition, another aim of this study is to focus on socioeconomic differences between clusters. Specifically the following hypotheses will be tested:

1. I expect three distinguishable clusters based on television watching and literacy activities, similar to the study of Phillips and Lonigan (2009). One cluster with high engagement in direct literacy activities engagement, high frequency in watching television and low engagement in indirect activities; one cluster with low engagement in all literacy activities and watching television; and one cluster with high engagement in all literacy activities and low frequency of watching television. There might also be a distinction between emphasis on direct and indirect literacy activities.
2. Based on the positive relation of HLE and emergent literacy skills, it is expected that the cluster with more literacy activities will have higher

emergent literacy outcomes. As a result, the three clusters would have the following characteristics: (a) high frequency in both direct and indirect literacy activities as well as in emergent literacy indicators and low frequency of watching television; (b) low frequency in both direct and indirect literacy activities as well as in emergent literacy indicators and high frequency of watching television; and (c) low frequency of indirect activities, high frequency of direct activities and emergent literacy indicators and high frequency of watching television

3. Negative correlations between watching television and the frequency of literacy activities and emergent literacy indicators are also expected.
4. Socioeconomic differences are expected to be found between the practices that parents adopt and, as a result, between clusters. Higher SES parents are expected to engage more frequently in literacy activities. As a result, clusters with high engagement in literacy activities are expected to be from high SES.

Method

Participants

Fifty eight Greek caregivers of children from 2 to 6 years of age took part in the study. The participants were recruited through personal contacts. The questionnaires were completed by either the mother or the father, as well as, in some cases by both, but not by others such as grandparents. The participants that hadn't completed crucial questions of the questionnaire were excluded. Moreover, those who had children older than 6 years old were excluded in order not to include children that already started school, and as a result formal instruction of writing and reading. The mean age of the children was 4.9 years (*SD*=1.0). 42 % of the sample were boys. The greater percent of the mothers (37%) had a Bachelor's

degree and 12 percent completed a Master's degree or higher education. Moreover, 10 percent followed vocational education and only 4 percent had stopped their studies at elementary school. The greater percent of fathers (32%) had a Bachelor's degree, followed by high school education (27%) and vocational education (20%). In addition, 12% had a Masters' degree or higher education, while 9 percent had only elementary education. The mean working hours of the mothers was 30 (*SD*= 19.3) hours and of the fathers 47.3 (*SD*= 15.1) hours. Almost half of the sample (47%) had a family income between 1.751€ and 3.200€ followed by families who had an income between 900€ and 1.250€ and families who had an income between 3.201€ and 4.201€. About twelve percent (11%) had an income between 1.251€ and 1.750€, while 10 percent had a income higher than 4.201€. Finally only 4% of the sample had family income less than 900€.

Measures

Home literacy environment questionnaire (HLE questionnaire). The HLE questionnaire was originally in English and translated in Greek by native Greek speakers. The questionnaire included questions about the demographic characteristics of the families, namely monthly family income, parents' education and parents' working hours.

In addition, the HLE included questions about parents' own literacy habits and their literacy activities together with the children. As far as parents' literacy habits are concerned, they were described by questions such as how often they buy a book, if they have a library subscription, if they read when their children are around, and an estimation of the number of books available in home. Questions concerning parental literacy habits were measured on a 5-point scale from 0 (*never*) to 4 (*more than once a week*), or in a 4 point-scale from 0 (*never*) to 3 (*very often*). Cronbach's alpha for parents' literacy habits was low (.50). Given that low values of Cronbach's alpha are not accepted, another measure concerning parents' literacy

habits, Adults' Author Checklist, was used in the analysis to measure parent's literacy habits, which is described in the following subsection below.

For describing the literacy activities, various questions were used concerning the indirect literacy activities such as reporting the number of products that are available at home (workbooks, educational games, educational video games and educational computer programs), frequency of library visits with the child, frequency of singing with the child, frequency of shared book reading in a typical week, and giving an estimation of the duration of shared reading in a typical week and of the number of children's books available at home. In addition some questions for describing direct literacy activities were used such as frequency of writing words with the child, frequency of reading letters with the child, frequency of writing letters with the child and frequency of teaching alphabet knowledge to the child. Questions concerning indirect literacy activities were measured on a 3-point scale from 0 (*almost never*) to 2 (*often*) and questions concerning direct literacy activities were measured on a 5-point scale from 0 (*not once in a month*) to 5 (*almost daily*). Cronbach's alpha for direct literacy activities was satisfactory (.86) while the same measure for indirect literacy activities had a low alpha reliability (.50). For the last aspect another measure was also used , Children's Storybooks Checklist, which is described in the following subsection below. Given that low values of Cronbach's alpha are not accepted, the checklist measure was used in the analysis to measure indirect literacy activities.

In addition, some questions concerning the emergent literacy of the children were included as indicators of emergent literacy, such as the frequency of writing letters or words, the frequency of the child writing letters of his or her name, the frequency of child's reading attempts and an estimation of the letters that the child already knows. Questions concerning the emergent literacy indicators were assessed on a 3-point scale from 0 (*never*) to 2 (*often*). Cronbach's alpha for emergent literacy indicators was satisfactory (.85).

Print exposure checklists. Additional information about HLE can be assessed with the use of a print exposure checklist. It is assumed that a person who reads frequently will be more familiar with literature and will recognize more book titles than a person who reads less (Mol & Bus, 2011). Two measures of book exposure were developed: Children's Storybook Checklist and Adults' Author Checklist to assess parents' exposure to adult literature. The construction of the two measures is described below.

Children's Storybook Checklist (CSC). For the assessment of children's book exposure, a checklist was designed similar to the procedure used in studies by Sénéchal, LeFevre, Hudson, and Lawson (1996). Titles of popular children's books were obtained through the bestselling lists of four major Greek bookstores and also through internet and more specific through one online Greek store (Perizitito.gr). Moreover, a list with the most famous storybooks according to Greek parents that was available on this website was used to compose the test. The final test consisted of 43 children's storybooks, 33 real and 10 foils in order to control guessing. Parents were instructed to select only story books that they recognized and to avoid guessing. For final coding the number of checked foils was subtracted from the number of checked real storybooks.

Adults' Author Checklist (AAC). An Adults' Author Checklist (AAC) was developed for the assessment of the parents' exposure to adult literature similar to the procedure of the study of Sénéchal et al. (1996). The bestsellers in adult fiction literature were the main focus and they were obtained through best selling lists of four major Greek bookstores and one online Greek store (perizitito.gr). The final list included Greek and foreign authors based on the best selling lists. The final test consisted of 45 names, 30 real authors and 15 foils, in order to control guessing. As in CSC, parents were instructed to select only those authors they recognized and to avoid guessing. Scoring procedure was similar as well, as to score this

measure the number of real authors and foil authors that have been recognized were calculated. Final scoring was consisted by subtracting the number of foils from the number of real authors.

Procedure

Parents received the questionnaires and the checklist via email or hand delivered in a hard copy. The parents were asked not to search at home any books when identifying the known authors or books in the checklists. They either completed them electronically or they answered with pen and pencil. The answered questionnaires were returned by email or by mail.

Data Analysis

The statistical analysis was conducted with SPSS 17.0. The numerical data were described by means, standard deviations, whereas the categorical data were described by the frequencies. From the outcomes of this analysis was found 3 missing values on mother's working hours, 4 on father's working hours, 1 in education of father and 6 in family income. Missing values were replaced with the mean in order not to lose valuable data of the sample. In order to examine if relationships between two categorical data exist, cross-tabs were used. The correlations of the numerical variables were measured with Pearson's correlation which is a measure that indicates linear relation between two variables.

Cluster analysis was conducted with the *K*-means method. Cluster analysis categorized the participants in clusters-groups where they have similar characteristics within a cluster but different characteristics among clusters. With Analysis of Variance it is examined if every variable is significant different between clusters. Cluster characteristics are illustrated by descriptive measures as described above. Moreover to examine if there are any

socioeconomic differences between clusters, Analysis of Variance was conducted for the numerical demographic variables. In addition, a confirmatory analysis was conducted with discriminant analysis to test the clusters results. The main purpose of discriminant analysis is to predict group membership, in this case cluster membership, based on a combination of interval variables, here the variables that had been used to predict cluster membership. The procedure begins with a smaller set of the data when the interval variables and cluster membership are considered as know. The end result of the procedure is a model that allows prediction of cluster members when only the interval variables are known. Finally, the comparison of cluster analysis and discriminant analysis was tested with contingency table, which is used to record and analyze the relation between two categorical variables, and Wilks's Lambda, in order to see the agreement of those two ways of analysis. High agreement will be an indicator of good classification, as two different ways of classification of the data will end up in the same results.

Results

Descriptive analysis

The descriptive characteristics of the variables are shown in Table 1. In general the caregivers read to their child more than 5 hours per week and owned on average more than 40 children's books. The caregiver reported engagement in letter-focused activities about once a week, on average. On the other hand, children were sometimes engaging in emergent literacy activities such as write words or their name and attempts to read. Moreover, they watched television more than 10 hours per week on average. In addition for caregivers' engagement in shared-reading activities the use of checklist showed that they recognized more than 6 authors from a total of 33, on average. The same value for the measure that corresponded to adult literature was recognizing more than 7 authors from a total of 30. Bivariate correlations

Table 1
Descriptive Characteristics of Home Literacy Environment

Characteristics	*M*	*SD*	*Range*
Caregiver reads letters with the child[a]	2.39	1.44	0-4
Caregiver writes letters with the child[a]	2.50	1.32	0-4
Caregiver writes words with the child[a]	2.20	1.41	0-4
Caregiver teaches alphabet to the child[a]	2.16	1.47	0-4
Child tries to read[b]	1.27	.74	0-2
Child writes some words[b]	1.41	.77	0-2
Child writes his/her name[b]	1.55	.75	0-2
AAC	7.38	5.96	0-30
CSC	6.30	4.56	0-33
Child watches television (hours per week)	12.94	8.62	0-39

Note. $N = 57$.
[a]Responses ranged from 0 (0 times in a moth) to 4(daily).
[b]Responses ranged from 0 (almost never) to 2(often).

for the 10 variables used in the cluster analysis are shown in Table 2. Significant correlation was found between all the variables of direct literacy activities frequency and between the variables that were used as indicators of emergent literacy. In addition, direct literacy activities were correlated with all the variables that were emergent literacy indicators. The measure that was used as the variable of indirect literacy activities (CSC), was correlated only with some variables of the emergent literacy indicators. Finally television watching was not significantly correlated with any variables.

Cluster analysis

Multiple K-means cluster analyses were conducted in order to conclude which model fitted the data. The three cluster solution appeared to best fit the data. As a result, the full sample was categorized into three clusters of 15, 32 and 10 participants respectively. The first cluster was characterized by high frequency in literacy activities but low on watching television, the second was characterized by low frequency of literacy activities and watching television and the third was characterized by low frequency of literacy activities and high frequency of watching television. This analysis included all the variables as shown in

Table 2
Bivariate Correlations Between Home Literacy Characteristics Included as Clustering Variables

Variables	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1.Caregiver reads letters with the child	-									
2.Caregiver writes letters with the child	.68**	-								
3.Caregiver writes words with the child	.71**	.70**	-							
4.Caregiver teaches alphabet to the child	.41**	.56**	.61**	-						
5.Child tries to read	.59**	.47**	.50**	.26	-					
6.Child writes some words	.67**	.57**	.65**	.36**	.58**	-				
7.Child writes his/her name	.58**	.58**	.61**	.34**	.55**	.83**	-			
8.AAC	.06	-.11	-.01	-.10	.11	.14	.20	-		
9.CSC	.10	.07	.01	.04	.28*	.13	.27*	.67**	-	
10.Child watches television(hours per week)	-.14	-.06	.01	.01	-.04	-.04	-.13	-.23	-.24	-

Note. N=57.
*p < .01. **p < .001.

Table 2 except for AAC variable, a measure of parents' literacy habits. Four variables indicated direct literacy activities, three indicated emergent literacy indicators and one variable indicated indirect literacy activities. In addition watching television was added as a variable that may be considered as a distracter of literacy activities or opportunities. However, one way analysis of variance indicated no significant mean difference for the majority of the variables (caregiver reads letters with the child, caregiver writes letters with the child, caregiver writes words with the child, caregiver teaches alphabet to the child, child tries to read, child writes some words) between the three or two cluster groups. Only for CBC variable, watching television and writing his/her name were found to differentiate between clusters, $F(2,54)= 65.81$, $p<.001$, $F(2,54)= 43.19$, $p<.001$ and $F(2,54)= 3.53$, $p<.05$, respectively.

Table 3
Cluster Characteristics in Aspects Included

Aspects	Cluster 1 HHL[a]	Cluster 2 LLL[a]	Cluster 3 LLH[a]
Direct Activities	High	Low	Low
Indirect Activities	High [b]	Low[b]	Low[b]
Emergent Literacy Indicators	High[c]	Low[c]	Low[c]
Watching Television	Low[d]	Low[d]	High[d]

[a] Clusters were named based on aspects that were found significant different at least in one variable of the aspect, between Clusters.
[b] Significant differences were found in CSC.
[c] Significant differences were found in child writing his/her name.
[d] Significant differences were found in watching television hours per week.

Cluster Descriptors

Inspection of the cluster profiles focusing on literacy activities, emergent literacy indicators and watching television lead to naming the clusters as being high or low based on one another and the sample average on these areas (see Table 3 and Table 4).

Cluster 1 (n=15), which represented 26.3% of the total sample, is the *"high-high-low" (HHL)* group characterized by high frequencies for both direct and indirect activities as well as for emergent literacy indicators (see Table 4 and Figure 1). This group had especially higher values on the variable write his/her name which is an emergent literacy indicator, in comparison with the other groups. As far as indirect literacy activities were concerned, this group had significant higher values when compared to the other groups and to the average value of the total sample. Moreover this group was characterized by the lowest number of television hours per week which was comparable to Cluster 2 but significantly lower than the Cluster 3. As far as secondary variables are concerned, such as descriptors of SES and parents' literacy habits, Cluster 1 and 2 were comparable in family income, maternal education and fathers' working hours. Moreover, although Cluster 3 seemed to be the group with the lowest SES, there weren't significant differences in SES variables between clusters. The only variable that was found to be significantly different between clusters was AAC (see Table 5). Cluster 1 had significant higher AAC values than the other Clusters.

Table 4
Means (Standard Deviations) for Home Literacy Characteristics for Clusters

	Cluster 1 "HHL"[a]	Cluster 2 "LLL"[b]	Cluster 3 "LLH"[c]
Caregiver reads letters with the child	$2.87_a(1.19)$	$2.26_b(1.53)$	$2.10_c(1.45)$
Caregiver writes letters with the child	$2.87_a(1.30)$	$2.36_b(1.28)$	$2.40_c(1.51)$
Caregiver writes words with the child	$2.67_a(1.35)$	$1.91_b(1.40)$	$2.40_c(1.43)$
Caregiver teaches alphabet to the child	$2.60_a(1.69)$	$1.94_b(1.41)$	$2.20_c(1.69)$
Child tries to read	$1.60_a(.51)$	$1.01_b(.78)$	$1.33_c(.82)$
Child writes some words	$1.73_a(.46)$	$1.28_b(.85)$	$1.34_c(.82)$
Child writes his/her name	$1.97_a(.12)$	$1.44_b(.80)$	$1.30_c(.95)$
CSC	$12.67_a(3.54)$	$4.17_b(1.88)$	$3.60_c(2.46)$
Child watches television (hours per week)	$9.47_a(4.93)$	$10.11_b(5.07)$	$27.20_c(6.73)$

Note. N = 57.
[a] *N*=15. [b] *N*=32. [c] *N*=10.

Cluster 2 (*n=32*), which represented 56.1% of the total sample, is the *"low-low-low" (LLL)* group and was characterized by having low frequencies on both direct and indirect activities as well as on emergent literacy indicators (see Table 4 and Figure 1) . Cluster 2, had comparable values on these aspects with Cluster 3 but significantly lower than Cluster 1 in indirect activities. In addition, this group had also low values on watching television hours per week which were comparable to Cluster 1 but significantly lower than Cluster 3. In external variables, Cluster 2 seemed to be comparable to Cluster 1 and higher than Cluster 3. This group had comparable AAC values with Cluster 3 but significantly lower than Cluster 1.

Finally, Cluster 3 (*n=10*), which represented 17.5% of the total sample, is the *"low-low-high" (LLH)* group (see Table 4 and Figure 1). This group had lower frequencies on some direct activities and emergent literacy indicators (caregiver reads letters to the child, caregiver writes letters with the child, child writes words and child writes his/her name) compared to both other clusters but in other direct activities and emergent literacy indicators (caregiver writes words with the child, caregiver teaches alphabet to the child, child tries to read) were found to have higher values than Cluster 2, but still lower than Cluster 1.

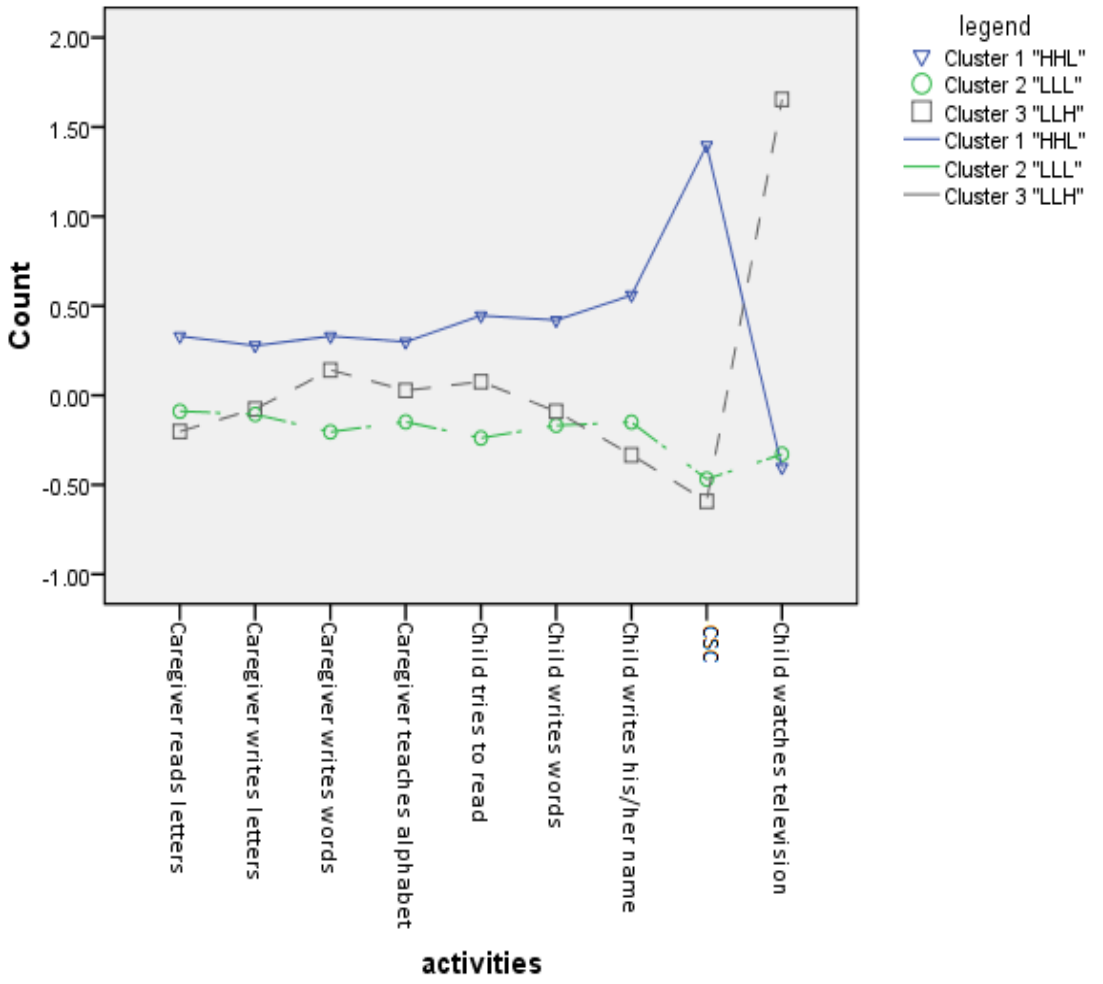

Figure 1. Line plot on performance on HLE for each of three clusters. *Note.* *n*=15 for Cluster 1, "high-high-low"; *n*=32 for Cluster 2, "low-low-low"; *n*=10 for Cluster 3, "low-low-high". All values were standardized within the total sample.

However none of these differences was found significant. Cluster 3 had comparable values with Cluster 2 in indirect activities but significantly lower values than Cluster 1. Finally, this group had significantly higher values in watching television hours per week than both other clusters. Cluster 3 seemed to have the lower SES characteristics, and significantly lower AAC values than Cluster 1.

Table 5
Descriptive Characteristics of Socioeconomic Variables for Clusters

Characteristics	Cluster 1	Cluster 2	Cluster 3
Family income	$3.90_a(1.03)$	$3.90_b(1.33)$	$3.48_c(.84)$
Mother's education	$3.53_a(.99)$	$3.53_b(1.02)$	$3.00_c(.67)$
Father's education	$3.40_a(1.30)$	$3.29_b(1.11)$	$2.62_c(.98)$
Mother's working hours	$28.53_a(14.77)$	$32.81_b(20.61)$	$23.10_c(17.72)$
Father's working hours	$46.75_a(13.52)$	$46.72_b(16.51)$	$50.00_c(8.74)$
AAC	$12.47_a(7.41)$	$6.04_b(4.26)$	$4.00_c(3.30)$

Note. N = 57.
[a] *N*=15. [b] *N*=32. [c] *N*=10.

Confirmatory Analysis

Wilks's lambda indicated two discriminant functions: the first strongly related to indirect literacy activities indicators, and the second strongly related to television watching and some direct literacy activities (see Table 6) . Less than 2 percent of the total data was misclassified with dicriminant analysis, and the misclassified data were only in Cluster 2 (see Table 7).

Discussion

The analysis resulted to a three-cluster solution that was distinguishable in indirect literacy activities, television watching and emergent literacy indicators. Both in correlations and clustering profiles, the results were consistent with those of several studies (Phillips & Lonigan, 2009; see also Bus et al., 1995; Sénéchal et al., 2008; Sénéchal et al., 1996). However, significant differences in direct activities engagement were not found between clusters, although expected.

Cluster 1 was characterized by significantly higher frequency in indirect activities and emergent literacy indicators than the other clusters and the average value of the total sample. However, when examining the variables that were significantly different, it is revealed that Cluster 1 had higher values in all variables when compared to all other clusters and the average of the total sample. For this reason it could be argued this cluster is the one with the

Table 6
Results of Discriminant Function Analysis to Predict Cluster Membership

	Correlation of predictor variables with discriminant functions		Clusters' group means		
Predictor Variable	1	2	Cluster 1	Cluster 2	Cluster 3
Caregiver reads letters with the child	-.02	-.22	2.87_a	2.25_b	2.11_c
Caregiver writes letters with the child	-.12	-.06	2.87_a	2.38_b	2.33_c
Caregiver writes words with the child	.29	.58	2.67_a	1.92_b	2.44_c
Caregiver teaches alphabet to the child	.16	-.02	2.6_a	1.97_b	2.11_c
Child tries to read	-.29	.19	1.60_a	1.10_b	1.33_c
Child writes some words	.31	.25	1.73_a	1.28_b	1.33_c
Child writes his/her name	-.02	-.46	1.97_a	1.45_b	1.22_c
CSC	.99	.45	12.67_a	4.13_b	3.67_c
Child watches television (hours per week)	-.43	.87	9.47_a	10.41_b	28.00_c
Canonical R	.88	.77			
Eigenvalue	3.49	1.47			
Group centroids					
Cluster 1	2.89	.61			
Cluster 2	-.67	-.95			
Cluster 3	-2.21	2.12			

Note. N = 57.
[a] *N*=15. [b] *N*=33. [c] *N*=9.

highest engagement in literacy activities in general. Moreover, this engagement is reflecting also to the emergent literacy indicators similar to the findings of other studies (Chaney, 1994; Sénéchal et al., 1996).

In contrast, Cluster 2 had significantly different engagement in literacy activities than Cluster 1. This cluster had lower engagement in both direct and indirect activities than Cluster 1 and in some variables than both other clusters. It is the cluster which was not highly engaged in any literacy activity and the one with the lowest values in two of the three emergent literacy indicators, when compared to other clusters and all indicators were lower

Table 7
Classification Rates from Discriminant Analysis of Cluster Variables

	Cluster 1	Cluster 2	Cluster 3
Correctly classified[a]	100%	97%	100%
Misclassified[a]		Cluster 3: 3%	

Note. N= 57.
[a] Calculated percentage was within clusters.

than the average values of the total sample. It could be argued that the low engagement in literacy activities reflected on the low values of emergent literacy indicators, reinforcing the hypothesis about the strong positive relation between HLE and emergent literacy as illustrated in Cluster 1 and to previous studies (Bus et al., 1995; Chaney, 1994; Sénéchal et al., 1996).

Finally, Cluster 3 was also a cluster with low engagement in literacy activities. However, when looking at the variables that weren't found significant, it can be assumed that in this cluster parents focused more on some direct activities such as writing words and teaching alphabet. It could be concluded that this engagement had as a result higher values in emergent literacy indicators than Cluster 2, which had comparable characteristics in literacy activities engagement with Cluster 3. These findings are consistent with those of Aram and Biron where joint writing was found more effective in print concepts than joint reading. In this case, emergent literacy indicators mainly concern print concepts, and this cluster had previous characteristics; parents seemed to focus more on direct teaching, which is described mainly by print concepts through the questionnaire, and achieve lower scores in CSC, which probably reveals the lack of joint reading. The findings are similar to those of Sénéchal et al., (1996) where parents' teaching contributed to children's written language skills.

In addition, watching television was expected to be negatively correlated to literacy activities, based on previous literature (Phillips & Lonigan, 2009). On the contrast there was no correlation found between watching television and any other variable, a fact which is consistent with other studies where watching television was not found significant correlated

to any measure (Patterson, 2002; Scarborough & Dobrich, 1994). On the other hand, watching television was significantly different among the clusters, in a way that was expected based on the study of Phillips and Lonigan (2009).

Moreover, demographic characteristics were expected to be different for various types of literacy activities engagement, based on several previous studies (Philips & Lonigan, 2009; see also Chaney, 1994; Van Steensel, 2006); they were expected to be positively correlated to literacy activities (Sénéchal et al., 1996). In this study, parents' literacy habits weren't found significantly correlated to any variable. In addition, demographic variables didn't significantly differentiate between clusters but parents' literacy activities were distinguishable. It is interesting, however, to thoroughly examine the relations between the demographic variables although considered not significant, since they reveal familiar trends. Cluster 1, whose members provided to children the most stimulating HLE, consisted of people with higher average income and mothers with higher educational status that Cluster 3 and the same as Cluster 2, a relation which is consistent with the study of Chaney (1994). It also had fathers with the highest educational status and parents with the highest literacy habits (Sénéchal et al, 1996). In contrast, although Cluster 2 is comparable with Cluster 1 as far as family income and mother's education are concerned, the working hours of the mothers are remarkably higher than in all other clusters. This should be the difference in SES characteristics, which could explain the different engagement in literacy activities. Although Cluster 2 had similar SES characteristics to Cluster 1, it had significantly different parents' literacy habits and engagement in literacy activities. It is the cluster which is not highly engaged in any activity. It could be assumed that the mothers' working hours result to less engagement to activities with the child. The third cluster was the cluster with the lowest SES, in terms of family income and in parents' education as well. Additionally, it is the cluster with the highest father's working hours and the lowest mother's working hours.

Limitations

The main limitation of the study is that although it is a study with a main focus on emergent literacy, it has no emergent literacy measure. All measures were used as indicators because there were parents' reports about a child behavior. Many tests concerning emergent literacy skills would be more appropriate and valid to measure children's emergent literacy skills. Another limitation of the study is that not only emergent literacy indicators, but all measures that have been used were self-reported. A children's Book Cover Recognition test, as designed and used in the study of Davidse et al, (2010) could be used in the future instead of a parent's checklist such as CSC. Moreover, the study had a rather small number of participants. It is possible that a larger number of participants would reveal relations that are now insignificant; namely, it could reveal significant differences in all variables among clusters or SES differences. In addition, this study did not calculate a sum variable for SES in order to categorize the participants to low, middle, and higher SES. This distinction was not clearly made and SES is described by many variables at a time. Finally, some questions in HLE questionnaire, although they referred to the same aspect, were measured in different scales (i.e., 3-point, 4-point, and 5-point scales). This may have caused the low Cronbach's Alpha value in two aspects (indirect literacy activities and parents literacy habits), since only those two aspects had different scales in questions concerning the same aspect, when the other aspects which had the same scaling, were found to have a good Cronbach's alpha.

Conclusions

In summary, the results enhance the findings of Philips and Lonigan (2009) about three different HLE environment profiles: (a) one that can be characterized as rich HLE as its members highly engage in literacy activities, and low watching television; (b) one with low engagement in literacy activities and watching television; and (c) one with low engagement

in literacy activities and high in watching television. Moreover, the relation between HLE and emergent literacy is supported both by correlation measures and significant differences among clusters. Although, television wasn't significantly correlated to any variable, it was significantly different between clusters. Socioeconomic status didn't arise any significant difference, yet parents' literacy habits were found significantly correlated to many variables and significantly different among clusters.

References

Aram, D. & Biron, S. (2004). Joint storybook reading and joint writing interventions among low SES preschoolers: differential contributions to early literacy. *Early Childhood Research Quarterly, 19,* 588-610. doi:10.1016/j.ecresq.2004.10.003

Bus, A. G., van IJzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research, 65,* 1–21. doi: 0.3102/00346543065001001

Chaney, C. (1994). Language development, metalinguistic awareness, and emergent literacy skills of 3-year-old children in relation to social class. *Applied Psycholinguistics, 15,* 371-394. doi: 10.1017/S0142716400004501

Davidse, N. J., de Long, M. T., Bus, A. G., Huijbregts, S. C. J. & Swaab, H. (2010). Cognitive and environmental predictors of early literacy skills. *Reading and Writing, 24,* 395-412. doi:10.1007/s11145-010-9233-3

Evans, M. A., & Shaw, D. (2008). Home grown for reading: Parental contributions to young children's emergent literacy and word recognition. *Canadian Psychology, 49(2),* 89-95. doi: 10.1037/0708-5591.49.2.89

Farver, J. M., Xu, Y., Eppe, S. & Lonigan, C. J. (2006). Home environments and young Latino children's school readiness. *Early Childhood Research Quarterly, 21,* 196–212. doi: 10.1016/j.ecresq.2006.04.008

Lonigan, J. C., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology, 36(5),* 596-613. doi: 10.1037//OOI2-1649.36.5.596

LeFevre, J. A., Polyzoi, E., Skwarchuk, S. L., Fast, L., & Sowinski, C. (2010). Do home numeracy and literacy practices of Greek and Canadian parents predict the numeracy skills of kindergarten children?. *International Journal of Early Years Education, 18*(1), 55-70. doi: 10.1080/09669761003693926

Mol, S. E. & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin, 137(2),* 267-196. doi: 10.1037/a0021890

Patterson, J. L. (2002). Relationships of expressive vocabulary to frequency of reading and television experience among bilingual toddlers. *Applied Psycholinguistics, 23,* 493-508. doi: 10.1017.S0142716402004010

Phillips, B. M., & Lonigan, C. J. (2009). Variations in the home literacy environment of preschool children: A cluster analytic approach. *Scientific Studies of Reading, 13*(2), 146-174. doi: 10.1080/10888430902769533

Roberts, J., Jurgens, J., & Burchinal, M. (2005). The role of home literacy practices in preschool children's language emergent literacy skills. *Journal of Speech, Language, and Hearing Research, 48,* 345-359. doi:10.1044/1092-4388(2005/024)

Scarborough, H. S. & Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Development Review, 14,* 245-302. doi:10.1006/drev.1994.1010

Sénéchal, M., Pagan, S. & Lever, R. (2008). Relations among the frequency of shared reading and 4-year-old children's vocabulary, morphological and syntax comprehension, and narrative skills. *Early Education and Development, 19(1),* 27-44. doi: 10.1080/10409280701838710

Sénéchal, M., LeFevre, J. A., Hudson, E. M., & Lawson, E. P. (1996). Knowledge of storybooks as a predictor of young children's vocabulary. *Journal of Educational Psychology, 88(3),* 520-536. doi:10.1037/0022-0663.88.3.520

Sénéchal, M., LeFevre, J. A., Thomas, E. M., & Daley, K. E. (1998). Differential effects of home literacy experiences on the development of oral and written language. *Reading Research Quarterly, 33(1),* 96-116. doi: 10.1598/RRQ.33.1.5

Sulzby, E., & Teale, W. (1996). Emergent literacy. In R. Barr, M.L. Kamil, P.B. Mosenthal, & P.D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 727-757). Mahway, NJ: Lawrence Erlbaum.

van Steensel, R. (2006). Relations between socio-cultural factors, the home literacy environment and children's literacy development in the first years of primary education. Journal of Research in Reading, 29(4), 367-382. doi: 10.1111/j.1467-9817.2006.00301.x

Whitehurst, G. J., & Lonigan, C. J. (1998) Child development and emergent literacy. *Child Development, 69(3),* 848-872. doi:10.1111/j.1467-8624.1998.tb06247.x