Investigating the reliability and validity of the Toddler Home Learning Environment (THLE) scale

James Hall1*, Kathy Sylva2, Pamela Sammons2, Rebecca Smees3, Maria Evangelou2, Teresa Smith4 and Jenny Goff2

1Southampton Education School, University of Southampton, Southampton, United Kingdom, 2Department of Education, University of Oxford, Oxford, United Kingdom, 3Department of Psychology, University of Sussex, Brighton, United Kingdom, 4Department of Social Policy and Intervention, University of Oxford, Oxford, United Kingdom

Home learning environments prior to school are well-known predictors of educational trajectories but research has neglected children aged under three. The new Toddler Home Learning Environment (THLE) scale is one response and this paper investigates its reliability and validity. The THLE is an adaptation of the Preschool HLE (PHLE) measure developed by the Effective Pre-School Primary and Secondary Education (EPPSE) investigation in the 1990s. The THLE was developed as part of the Evaluation of Children’s Centers in England (ECCE) investigation that followed a sample of 2,608 of families from 14 to 38 months. The THLE was administered at 14 months, the PHLE at 38. The 8-item THLE evidences internal consistency via statistical reliability coefficients and Confirmatory Factor Analysis plus measurement validity via statistically significant and research-appropriate associations with the PHLE, three measures of child development, and child and parent demographics. This paper moves the HLE literature forward with a new parental self-report scale of the HLE that is for use with toddlers.

Keywords: home learning environment, parent-child interaction, toddlerhood and early relationships, validity and reliability analyses, toddlers

INTRODUCTION

Informed by social-constructivist theories (e.g. Bruner 1978; Vygotsky, 1978) and ecological theories (e.g. Bronfenbrenner and Morris, 2006) of learning and child development, the Home Learning Environment (HLE) that surrounds children and young people has been subject to extensive study and measurement (see Lehrl et al., 2020). Although there is a trend to differentiate literacy and numeracy aspects of the HLE (e.g. Niklas et al., 2020), significant cross developmental domain effects are commonly found (e.g. Niklas and Schneider, 2017) and these support the continuing specification of single-scale HLE measures (especially for younger age groups). However and despite its extensive study, the international literature base on the HLE still retains systematic knowledge gaps (e.g. Silinskas et al., 2020), one of which concerns the HLE for children aged under age three years (a relatively infrequently studied period; Burghardt et al., 2020) -- a gap to which this paper responds.

Given that research on the HLE for children of any age stresses the importance of adult-child interactions (e.g. Crampton and Hall, 2017; Lehrl et al., 2020), consideration of parenting practices linked to child development in the under-threes can foster understanding of the toddler HLE. These parenting practices include (but are not limited to):
• Shared reading activities that promote both emergent literacy skills (Rodriguez and Tamis-LeMonda, 2011) and language development via high-quality verbal interactions (Hayes and Berthelsen, 2020);
• Activities that regulate a child’s arousal, distress, and sensory stimulation (eventually internalized as self-regulation; Posner et al., 2014);
• Parents (typically) acting as their child’s first play partners which benefit the development of (cross-contributing) motor, cognitive, and social skills (e.g. Dinkel and Snyder, 2020).

A well-established and robust international evidence base exists (and continues to grow; e.g Melhuish et al., 2008; Bonci et al., 2011; Romeo et al., 2018) that attests to the long-term and sizable positive impacts of the HLE upon educational trajectories, educational equity, and long-term developmental outcomes (e.g. Jeynes, 2005; Son and Morrison, 2010; Sammons et al., 2015b; Shuey and Kankaras, 2018). The HLE in the preschool years is particularly important because it can have: 1. Effects on attainment through to adolescence (e.g. Cunningham and Stanovich, 1997); 2. Effects on attainment that are above and beyond those associated with social disadvantage (e.g. Flouri and Buchanan, 2004); and 3. The potential to partially attenuate the detrimental effects of social disadvantage on developmental and educational outcomes (e.g. Ramey and Ramey, 2004).

A large volume of evidence concerning the long-term impacts of the HLE -- and of the preschool HLE in particular -- comes from the Effective Pre-School, Primary and Secondary Education (EPPSE) project (see Sammons et al., 2004; Melhuish et al., 2008; Sammons et al., 2015b; Toth et al., 2020). This was a prospective nationally representative longitudinal study that took place in England between 1997 and 2014, followed a sample of 3,000 + children from mean age three years through to age 16, and was the first large-scale United Kingdom investigation to focus upon the effectiveness of early years education. While it is outside the remit of this paper to review all of EPPSE’s findings (for details see Sylva et al., 2010), it is important to note that the tool developed by EPPSE to measure the HLE in the preschool years, here termed the Preschool Home Learning Environment (PHLE) scale, features in EPPSE publications that reveal long-term effects (e.g. Baker et al., 2014) and all three of the types of HLE effects documented above (e.g. Sammons et al., 2014; Sammons et al., 2015b; Toth et al., 2020). Partially because of such effects, EPPE’s PHLE scale has since featured in several subsequent large-scale and high-profile prospective longitudinal studies of child development and education including the Millennium Cohort Study (MCS; Dearden et al., 2011; de la Rochebrochard, 2012) and the Study of Early Education and Development (SEED; Melhuish et al., 2017). It has also influenced the development of measures of the HLE in other countries including the BiKS longitudinal study in Germany (Anders et al., 2013; Sammons and Anders, 2015).

However, despite growing international evidence documenting the mid and longer term effects of the HLE experienced by children from age three years, much less is known about the long-term effects from the HLE experienced by very young children under age three (Dodici et al., 2003; Burghardt et al., 2020) because it has been less frequently assessed (e.g. Elardo and Bradley, 1981) as compared to those aged three years and up. While it is beyond the scope of this paper to speculate on the reasons for this neglect, it is important to recognize that this initial period of life (encompassing infancy and toddlerhood) is characterized by rapid growth and change, particularly for language development (e.g. Rodriguez and Tamis-LeMonda, 2011), and equally rapid changes in adult-child interactions and activities that are developmentally appropriate (e.g. Brophy-Herb et al., 2018). Thus, understanding the effects of the HLE experienced during the first three years of life is no less important than understanding the effects from the HLE at later ages. For quantitative research to document these effects, an assessment of the HLE for toddlers is required that demonstrates measurement validity and reliability.

This paper responds to the (comparative) gap in knowledge regarding the effects of the HLE for the under-threes by investigating the reliability and validity of a parental self-report assessment tool (an adaption of EPPSE’s PHLE) that measures the activities that take place between adults and toddlers in the homes that support their learning. A demonstrably reliable and valid measure of toddlers’ HLE would help researchers, practitioners and policy makers to better understand the impacts of the activities that adults engage in with their toddlers that support their learning and enhance outcomes. In turn, this improved understanding has the potential to shape policy and practice by increasing our understanding of the drivers of child development, educational progress, and ultimately supporting greater educational equity.

The research question addressed by this paper is, Is the Toddler Home Learning Environment (THLE) scale a reliable and valid measure of the activities and resources that support children’s early learning at home?

This research question was answered via statistical analysis of data from the nationally-representative Evaluation of Children’s Centers in England (ECCE) project (2009–2015; details below) and through appraisal of several aspects of scale reliability and validity:

1. For a reliable THLE scale: First, a lack of homogeneity, floor, and ceiling effects in parent responses to THLE items. Second, statistically satisfactory evidence supporting the specification of a single THLE scale and the consistency of the THLE scale’s items in indicating this scale (details within the below Analytic Approach Section).
2. For a valid THLE scale: Satisfactory statistical evidence of measurement validity via examination of various combinations of criterion validity (via concurrent and predictive validity) with construct validity (via convergent and discriminant validity). First, predictive convergent validity was investigated via estimation of the association between the THLE and the PHLE. Second, the concurrent (and predicted) convergent and discriminant validity of the THLE was investigated via estimation of its association with other measures taken at the baseline (toddler) assessment point of the ECCE study. Third, further evidence of
measurement validity across both the criterion and construct domains was obtained by comparing how HLE associations with other measures changed (or remained constant) across the toddler and preschool period.

Together, these measurement validation analyses extend published ECCE findings that have already shown the THLE scale to significantly and positively predict three measures of child development during the preschool period (over and above a range of background measures including child age, gender, and health, plus family socioeconomic status; see Sammons et al., 2015a): verbal cognitive abilities, non-verbal cognitive abilities, and prosocial behavior. This latter association being particularly important because it reinforces findings elsewhere that stress the importance of the HLE for socioemotional competency (e.g. Wirth et al., 2020).

MATERIALS AND METHODS

Design

The Evaluation of Children’s Centers in England (ECCE) project was a prospective longitudinal study that followed a sample of 2,608 children (and their families) from mean age 14 months to mean age 38 months—all of whom were registered users at one of 117 Sure Start Children’s Centers (SSCCs; see Sylva et al., 2015). The ECCE project used a research design with five strands to meet five project objectives (one strand per objective): 1. To reveal models of SSCC leadership structure; 2. To identify patterns in the types of services that were used by families; 3. To identify common patterns of services offered by SSCCCs; 4. To identify the impact of SSCCs on child, mother, and family outcomes; and 5 (based on results from Strands 1–3); 5. To identify the cost-effectiveness of SSCCs (based on results from Strands 2 and 4).

Informed by the findings of previous studies carried out in the early years (particularly EPPSE, Sylva et al., 2010; and the National Evaluation of Sure Start, Belsky et al., 2006), positive changes to HLEs were one of the family outcomes upon which SSCCs were hypothesized to impact. As such, HLEs were measured at both study outset (via the ECCE developed Toddler Home Learning Environment scale; THLE) and again at mean child age 38 months (via use of the EPPSE developed PHLE scale).

Sample

There are two samples of families considered in this paper: A sample of 5,717 who took part in baseline assessments in 2012 (when their children were mean age 14 months) and a sub-sample of 2,608 who were followed up in 2014 because their SSCCs were taking part in a parallel longitudinal study (see Maisey et al., 2015). This systematic selection of families for follow-up resulted in samples with differing demographic characteristics. While the proportions of male and female children did not significantly differ between the 2,608 and the remaining 3,109 families (X^2(1, n = 5,717) = 2.07, p = 0.151), nor did the mental health of mothers at baseline (t (5319.87) = 1.29, p = 0.196), this was not the case for other demographic characteristics. Instead, the longitudinal sample of 2,608 families featured significantly greater numbers of: White British families and fewer families from Pakastani and “Mixed Race” backgrounds (X^2(8, n = 5,708) = 28.31, p < 0.001), mothers who held a higher level of qualification (X^2(6, n = 5,683) = 196.41, p < 0.001), and households with higher average incomes (U = 2,776,755.50, n = 5,199, p < 0.001).

Because the baseline sample and longitudinal sample differed from one another on a number of demographic characteristics, the two samples were put to different purposes within this paper. The baseline data from the 5,717 families were used within statistical appraisal of the reliability of the THLE scale (Research Question 1; making use of all THLE data obtained from the 5,717 families) while the longitudinal data from the 2,608 families were used in the appraisal of measurement validity (Research Question 2) as this relied upon comparing THLE scores to PHLE scores and the PHLE scores were only sought from the 2,608. Full details on these analyses are provided in the Analytic Approach Section below.

In terms of sample to population representativeness, the ECCE project followed a sample of SSCCs that were representative of Phase 1 and 2 SSCCs in England between 2009 and 2014 (see Tanner et al., 2012) and the 2,608 families who were followed over time were all registered users of one of these centers. The result is a sample of families who are broadly representative of those families who used Phase 1 and 2 SSCCs in England between 2012 and 2014 (Sammons et al., 2015a).

Measurement of the sampled families’ PHLE was taken on average 24 months after the baseline assessments and was only carried out within the purposively selected longitudinal sample of 2,608 families. At both measurement points (mean child ages 14 and 38 months), assessments of HLE and measures describing children and families, were carried out by a trained team of fieldworkers who visited the home of each child (Maisey et al., 2013; Maisey et al., 2015). Detailed information regarding the characteristics of the sampled SSCCs is found in Goff et al. (2013), Evangelou et al. (2014) and Sylva et al. (2015), information regarding the baseline sample of children and families in Maisey et al. (2013); as can details of the project’s research ethics), and full information on the longitudinal sample of children and families who participated in the impact analyses is given by Sammons et al. (2015a).

Measures

The Toddler Home Learning Environment (THLE) scale was developed by the ECCE team as an adaptation of EPPSE’s PHLE scale. The THLE measure was designed to serve as a baseline assessment (at mean child age 14 months) of the various developmentally appropriate activities that toddlers took part in alongside their adult caregivers. Primary caregivers (PCGs—of whom 96% were mothers; see Sammons et al., 2015a) self-reported the frequency with which their toddler engaged in seven of these activities in face-to-face interviews where their responses were recorded using 7-point frequency rating scales (coded 1 to 7). Accompanying these seven frequency questions was an extra question that asked PCGs to report the number of books in the home that were for the toddler (again using a 7-point rating scale). This non-frequency question was developed in recognition of the fact that the opportunity to engage in more frequent adult-toddler activities related to literacy acquisition is
TABLE 1 | Wording, response options, and descriptive statistics for the eight items developed for potential inclusion in the Toddler Home Learning Environment (THLE) scale, the two items relating to toddler television watching, and the seven items included in the Preschool Home Learning Environment (PhLE) scale.

| Toddler home learning environment (THLE) items | Parental self-report response options |
|----------------------------------------------|--------------------------------------|
| 1. How often does someone at home take (child) out of the house, for example visiting family or friends, or going to the park? (n = 5,714; Median = 6; Inter-Quartile range [IQR] = 3) | 1. Very rarely  
2. Once a week  
3. Twice a week  
4. Three times a week  
5. Four times a week  
6. Five or six times a week  
7. Every day or more than once a day |
| 2. How often does someone at home draw (child)’s attention to the names of things during their day-to-day activities? (n = 5,712; Median = 7; IQR = 0) | 1. Never/Not yet  
2. Have done this once or twice  
3. Less than once a week  
4. Once a week  
5. Several times a week  
6. Once a day  
7. More than once a day |
| 3. How often does someone use blocks or shape sorting toys with (child)? (n = 5,710; Median = 6; IQR = 2) | 1. No books  
2. 1–2 books  
3. 3–4 books  
4. 5–10 books  
5. 11–15 books  
6. 16–20 books  
7. 21 or more books |
| 4. How often does someone at home talk about, or try to teach (child) the names of colors or shapes? (n = 5,711; Median = 6; IQR = 2) | 1. Never/Not yet  
2. Have done this once or twice  
3. Less than once a week  
4. Once a week  
5. Several times a week  
6. Once a day  
7. More than once a day |
| 5. How often does someone at home sing songs or nursery rhymes to or with (child)? (n = 5,712; Median = 7; IQR = 1) | 1. Never/Not yet  
2. Have done this once or twice  
3. Less than once a week  
4. Once a week  
5. Several times a week  
6. Once a day  
7. More than once a day |
| 6. How often does (child) get a chance to play in a messy way, for example using playdough, paints, or sand? (n = 5,710; Median = 2; IQR = 3) | 1. Never/Not yet  
2. Have done this once or twice  
3. Less than once a week  
4. Once a week  
5. Several times a week  
6. Once a day  
7. More than once a day |
| 7. Although (child) is very young, some children do enjoy being read to or handling books designed for babies. How often does someone at home read to (child)? (n = 5,713; Median = 6; IQR = 2) | 1. No books  
2. 1–2 books  
3. 3–4 books  
4. 5–10 books  
5. 11–15 books  
6. 16–20 books  
7. 21 or more books |
| 8. How many books written for babies or toddlers does (child) have? (n = 5,712; Median = 5; IQR = 3) | 1. No books  
2. 1–2 books  
3. 3–4 books  
4. 5–10 books  
5. 11–15 books  
6. 16–20 books  
7. 21 or more books |

| Toddler television watching items |  |
|---------------------------------|---|
| A. How much time does the child spend watching TV/DVDs on their own while someone else is busy in the household? (n = 5,709; Median = 2; IQR = 2) | 1. None  
2. Fewer than 30 min per day  
3. 30 min – 1 h per day  
4. 1–2 h per day  
5. 2–3 h per day  
6. 3–4 h per day  
7. More than 4 h per day |
| B. How much time does the child spend watching TV/DVDs together with someone at home? (n = 5,708; Median = 1; IQR = 1) | 1. Never/Not yet  
2. Have done this once or twice  
3. Less than once a week  
4. Once a week  
5. Several times a week  
6. Once a day  
7. More than once a day |

| Preschool home learning environment (PhLE) items |  |
|-----------------------------------------------|---|
| 1. How often does someone at home read to (child)? (n = 2,576; Median = 4; IQR = 1) | 1. Never  
2. Have done this once or twice  
3. Less than once a week  
4. Once a week  
5. Several times a week  
6. Once a day  
7. More than once a day |
| 2. How often does someone at home take (child) to the library? (n = 1,432; Median = 2; IQR = 1) | 1. Never  
2. Have done this once or twice  
3. Less than once a week  
4. Once a week  
5. Several times a week  
6. Once a day  
7. More than once a day |
| 3. How often does (child) play with letters at home? (n = 2,193; Median = 4; IQR = 5) | 1. Never  
2. Have done this once or twice  
3. Less than once a week  
4. Once a week  
5. Several times a week  
6. Once a day  
7. More than once a day |
| 4. How often does someone at home help (child) to learn the ABC or alphabet? (n = 2,326; Median = 4; IQR = 5) | 1. Never  
2. Have done this once or twice  
3. Less than once a week  
4. Once a week  
5. Several times a week  
6. Once a day  
7. More than once a day |
| 5. How often does someone at home try to teach (child) numbers or counting? (n = 2,571; Median = 7; IQR = 3) | 1. Never  
2. Have done this once or twice  
3. Less than once a week  
4. Once a week  
5. Several times a week  
6. Once a day  
7. More than once a day |
| 6. How often does someone teach (child) songs, poems or nursery rhymes? (n = 2,513; Median = 7; IQR = 3) | 1. Never  
2. Have done this once or twice  
3. Less than once a week  
4. Once a week  
5. Several times a week  
6. Once a day  
7. More than once a day |
| 7. How often does (child) paint or draw at home? (n = 2,536; Median = 4; IQR = 4) | 1. Never  
2. Have done this once or twice  
3. Less than once a week  
4. Once a week  
5. Several times a week  
6. Once a day  
7. More than once a day |

IQR: Inter-Quartile Range.

Constrained by the availability of developmentally appropriate materials (Bradley and Caldwell, 1995), and that the availability of these materials will be influenced, at least in part, by household income (e.g. De Bondt et al., 2020).

During the development of the THLE, two items were devised that focused not on the THLE (on activities concerning adult-toddler interaction) but instead on toddler television watching as a kind of ‘displacement activity’ (e.g. Dore et al., 2020) from the interactions that promote learning. These two items were developed alongside the THLE items in order to equip the ECCE study with the ability to describe other things that toddlers may have been doing other than engaging in interactions linked to learning and development. The full text of the items developed to reflect the THLE and the items developed to capture toddler television watching (plus the response options to these items) is shown in Table 1 with descriptive statistics. Statistical results that informed the development of the THLE scale are reported in the Results.
The Preschool Home Learning Environment (PHLE) scale is the same (primary caregiver reported) measure of HLE originally developed by the EPPSE project (see Sammons et al., 2004; Melhuish et al., 2008; Sylva et al., 2010). The PHLE measure is an index created from the summation of seven items that record the frequency with which seven adult-child shared play and learning activities are carried out with preschool-aged children (3–5 years) using 7-point frequency scales with responses that range from 1 to 7. The resulting PHLE measure that was constructed and used in the ECCE investigation had scores ranging from 7 to 49 (n = 2,604; mean = 30.59; standard deviation = 9.08). The items contributing to the PHLE, caregivers’ response options to these items, and the median response options are shown in Table 1. The statistical evaluation of the reliability and validity of the THLE scale that is reported in this paper was enriched through use of the additional data that were gathered by the ECCE study for the same sample of children and families who had their HLEs measured in the toddler and preschool years. Three sets of measures (child, primary caregiver, home environment) were included and used in different, though mutually informative, measures (child, primary caregiver, home environment) measured in the toddler and preschool years. Three sets of measures (child, primary caregiver, home environment) were included and used in different, though mutually informative, measures (child, primary caregiver, home environment) were included and used in different, though mutually informative, measures (child, primary caregiver, home environment). The statistical evaluation of the reliability and validity of the PHLE measure is presented in Table 2.

Table 2 presents summary descriptive statistics for the measures that were included in this study. The highest qualification (academic or vocational) held by each Primary caregiver (PCG) in a household was measured across individuals by comparing each reported qualification to its equivalent National Vocational Qualification (NVQ) level (where a higher NVQ level indicates a higher level of achieved qualification). Within this system an NVQ Level 1 captures qualifications equivalent to those from compulsory age 16 national assessments, and an NVQ Level 5 captures qualifications equivalent to (and including) university degrees (Lester, 2018). For readers unfamiliar with this practice of standardizing qualifications to NVQ levels, this is a method that has long been routine in United Kingdom educational practice, policy, and research (e.g. Dearden et al., 2002; Gayle et al., 2015).

Looking at the other measures presented within Table 2 that may require further explanation, within the home tenure measure, the “rent free” category captured households who were legally living, at zero cost, in a property that was owned by someone else (e.g. a friend or a relative). This category does not include households whose accommodation was lived in illegally, accommodation that was financially contributed to by the United Kingdom State due to low household income (this would still be rented accommodation), or accommodation that was owned (with or without a mortgage) by one or more members of the household. PCG mental health/well-being was measured via the General Health Questionnaire (GHQ; Goldberg and Williams, 1988) when children were both 14 (‘baseline’/toddler; (Cronbach’s $a = 0.88$) and 38 months.

### Table 2: Descriptive statistics for the additional measures included in this study.

| Measure                                           | Categories                              | n   | % or mean ± sd |
|---------------------------------------------------|-----------------------------------------|-----|----------------|
| Child gender (binary; male coded high)            |                                         | 5,717 | 51.1%          |
| Child age at “baseline” (months)                  |                                         | 5,708 | 14.12 ± 1.76   |
| Child age at “outcome” (months)                   |                                         | 2,608 | 38.14 ± 2.82   |
| Mother age at “baseline” (years)                  |                                         | 5,677 | 30.39 ± 5.89   |
| Baseline highest PCG qualification (n = 5,683)    | No qualification*                       | 520  | 9.1%           |
|                                                   | NVQ level 1                             | 408  | 7.1%           |
|                                                   | NVQ level 2                             | 1,366 | 23.9%          |
|                                                   | NVQ level 3                             | 1,001 | 17.5%          |
|                                                   | NVQ level 4                             | 1,541 | 27.0%          |
|                                                   | NVQ level 5                             | 638  | 11.2%          |
|                                                   | “Other”                                 | 209  | 3.7%           |
| Baseline home tenure (n = 5,710)                  | Rent free*                             | 230  | 4.0%           |
|                                                   | Rent it                                 | 2,737 | 47.9%          |
|                                                   | Shared ownership                        | 76   | 1.3%           |
|                                                   | Mortgage                                | 2,400 | 42.0%          |
|                                                   | Own it outright                         | 267  | 4.7%           |
| Baseline household income (n = 5,199)             | < £4,999                                | 191  | 3.3%           |
|                                                   | £5,000–£5,999                           | 757  | 13.2%          |
|                                                   | £10,000–£19,999                         | 1,198 | 21.0%          |
|                                                   | £20,000–£29,999                         | 957  | 16.7%          |
|                                                   | £30,000–£39,999                         | 761  | 13.3%          |
|                                                   | £40,000–£49,999                         | 589  | 10.3%          |
|                                                   | > £50,000*                              | 748  | 13.0%          |
| Toddler PCG: mental health/well-being             |                                         | 5,330 | 22.56 ± 5.59   |
| Toddler PCG: parenting stress (“parental distress” subscale) |          | 5,455 | 26.30 ± 8.09   |
| Preschool PCG: mental health/well-being           |                                         | 2,449 | 22.15 ± 4.98   |
| Preschool PCG: parenting stress (“parental distress” subscale) |          | 2,553 | 25.86 ± 7.75   |

sd, standard deviation; *Reference category when the measure was subsequently represented with dummy-coded binary measures; sd, standard deviation; PCG, Primary Caregiver; NVQ, National Vocational Qualification.
Stress Index (PSI; Abidin, 1995; alpha at 14 m = 0.95 and at 38 m = 0.92).

Analytic Approach

The statistical appraisal of the reliability and validity of the THLE items and the resulting THLE scale was grounded in Classical Test Theory (CTT) and was undertaken in two stages using a combination of SPSS version 24 (IBM Corporation, 2016), JASP version 0.10.2 (JASP Team, 2019), and Mplus version 7.4 (Muthén and Muthén, 2015). First, reliability was investigated using a mix of item-level analysis, statistical estimators, and confirmatory factor analyses that provided complementary lenses through which to examine the internal consistency of different combinations of the ten THLE and toddler television watching items within the baseline sample of 5,717 families. Second, the measurement validity of the THLE was appraised through statistical analyses of different aspects of criterion and construct validity (and via comparison with PHLE scores) within the longitudinal sample of 2,608 families.

To investigate the reliability of the THLE items, a range of statistical analyses were undertaken of the eight THLE items and two toddler television-viewing-items to determine their ability to serve as reliable indicators of an underlying latent THLE. First, Spearman correlation coefficients (rs) were used to consider the degree of shared response between each pair of items. Second, the Cronbach’s alpha statistic was then used to appraise each potential item’s contribution to an overall THLE scale. However, this analysis of the internal consistency of the items also responded to contemporary criticisms of the trustworthiness of Cronbach’s alpha for this purpose (e.g. Sijtsma, 2009; Dunn et al., 2014). As a result, internal consistency was also appraised through estimation of Guttmann’s (1945) five other lambda statistics (λ1, λ2, plus λ4 through λ6; Cronbach’s α = λ3; see Sijtsma, 2009) and McDonald’s omega (ω; McDonald, 1999). We also acknowledged the common practice of accepting the critical cut-off value for an acceptable alpha of 0.70 (Nunnally and Bernstein, 1994) but are aware of the limitation of accepting 0.70, or any other value, as a binary cut-off (Dunn et al., 2014). Finally, a Confirmatory Factor Analysis was undertaken of those items that the prior procedures had indicated demonstrated sufficient consistency to evidence an underlying (THLE) scale -- this to test the appropriateness of the unidimensional latent (THLE) construct (via model fit with data) that the prior statistical analyses pointed to.

Following from the reliability analyses, the association between the resulting THLE scale and the varying ages of the toddlers within the ECCE sample was then assessed because of the consequences that the fast pace of development at this age might have for the forms of adult-child shared play and learning activities that are developmentally appropriate. The result of this analysis was the development of an age-adjusted THLE scale (adj. THLE) that, along with a THLE scale that was based upon summation of items, was taken forward into the second stage of the analysis in this paper.

With a THLE scale developed from indicative items and demonstrating internal consistency, Stage two of the analysis then considered the measurement validity of the THLE through statistical procedures that made use of other measures collected by the ECCE researchers and that built on past published work (Sammons et al., 2015a) showing the THLE as a significant predictor of child cognition and behavior at mean age 38 months (see the Introduction). The result was a mix of statistical assessments that combined aspects of criterion validity (via concurrent and predictive validity) with aspects of construct validity (via convergent and discriminant validity). This began with consideration of the association shared between the two forms of the THLE and then progressed to consideration of the association between the THLE and the PHLE to evaluate the new scale’s predictive convergent validity over time (from average age 14 months to average age 38 months).

Second, the associations between the THLE (in both simple and age-adjusted forms) and other measures taken at the baseline (toddler) assessment point of the ECCE study were then estimated in order to evaluate the THLE’s concurrent (and predicted) convergent and discriminant validity. These associations were also compared to equivalent associations for the PHLE in order to demonstrate how HLEs prior to school entry can be stratified both consistently and uniquely when comparing the toddler years to the preschool years for the same sample. Multilevel statistical regression models that accounted for the nesting of families within children’s centers (a key feature of the ECCE sampling strategy) were used for these analyses. If the association between the THLE and demographics were found to be similar to the associations between the PHLE and various key demographics, then this would provide further evidence of measurement validity across both the criterion and construct domains. To facilitate this comparison between the THLE and PHLE, multilevel effect sizes were calculated following the approach used by Elliot and Sammons (2004).

An examination of the proportions of variance in the THLE and PHLE scores that were attributable to differences between SSCCs (rather between families; via Intra Class Correlations, ICCs) provided further support for the use of multilevel statistical regression models. Hox (2010) describes ICC values of 5, 10, and 15% as showing a small, medium and high effect of a nested sample design. Here, the THLE returned an ICC of 15% (SSCC variance = 8.10, p < 0.001), the age-adjusted THLE an ICC of 15% (SSCC variance = 0.13, p < 0.001), and the PHLE an ICC of 5% (SSCC variance = 4.20, p < 0.001). Thus, the nested design of the ECCE study resulted in non-trivial variation in THLE scores due to differences between SSCCs—variation captured by the multilevel regression models. Further, using this modeling approach for all of the HLE variables provided a consistent basis of comparison for demographic correlates across the THLE and PHLE.

RESULTS

Reliability

Item-Level Analysis

Further considering the descriptive statistics presented in Table 1, the medians and distributions of the THLE items and toddler television watching items demonstrated that these items were...
TABLE 3 | Bivariate associations between the eight items measuring the Toddler Home Learning Environment and two items measuring toddlers’ television watching environments (Spearman’s Correlations, rs)

| THLE and television watching items | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | A   | B   |
|-----------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1. How often does someone at home take (child) out of the house, for example visiting family or friends, or going to the park? | rs  | 1   |     |     |     |     |     |     |     |     |
| 2. How often does someone at home draw (child)’s attention to the names of things during their day-to-day activities? | rs  | 0.20*** | 1   |     |     |     |     |     |     |     |
| 3. How often does someone use blocks or shape sorting toys with (child)? | rs  | 0.17*** | 0.31*** | 1   |     |     |     |     |     |     |
| 4. How often does someone at home talk about, or try to teach (child) the names of colors or shapes? | rs  | 0.16** | 0.33*** | 0.49*** | 1   |     |     |     |     |     |
| 5. How often does someone at home sing songs or nursery rhymes to or with (child)? | rs  | 0.14** | 0.33*** | 0.29** | 0.34*** | 1   |     |     |     |     |
| 6. How often does (child) get a chance to play in a messy way, for example using playdough, paints, or sand? | rs  | 0.11*** | 0.12** | 0.16** | 0.18*** | 0.12*** | 1   |     |     |     |
| 7. Although (child) is very young, some children do enjoy being read to or handling books designed for babies. How often does someone at home read to (child)? | rs  | 0.22*** | 0.29*** | 0.27*** | 0.34*** | 0.33*** | 0.22*** | 1   |     |     |
| 8. How many books written for babies or toddlers does (child) have? | rs  | 0.22*** | 0.25*** | 0.19*** | 0.23*** | 0.20*** | 0.25*** | 0.45*** | 1   |     |
| A. How much time the child spends watching TV on their own, while someone else is busy in the household? | rs  | 11*** | -0.12*** | -0.07*** | -0.07*** | -0.10*** | -0.04*** | -0.10*** | -0.11*** | 1   |
| B. How much time the child spends watching TV together with someone at home? | rs  | 11*** | -0.07*** | 0.00 | 0.02 | -0.05*** | -0.02 | -0.05*** | -0.10*** | 0.47*** | 1   |

* p < 0.05; ** p < 0.01; *** p < 0.001.

differentially sensitive to the frequency with which different activities took place in the home. None of the ten items had blank responses for any of their seven response options (the least frequent response, from 29 respondents, was, “less than once a week” regarding how often a child’s attention was drawn to the names of things) and these items differed from one another as regards which of the activities were more and less common for toddlers to experience. The most frequent activities included the child having their attention focused on the names of objects during day-to-day activities, and exposure to songs or nursery rhymes (shared or otherwise). The average child was reported by caregivers as experiencing both of these activities more often than once a day. Slightly less frequently experienced (once a day for the average toddler) were someone using blocks or shape sorting toys with the child, teaching them the names of colors and/or shapes, and someone reading to the child. By contrast, the average toddler engaged in messy play much less frequently—on average reported only once or twice (at all); although with a high degree of variation (e.g. 8.1% of caregivers reported their child engaged in messy play every day or more often). Caregivers reported having an average of 11–15 books written for babies or toddlers at home—again there was substantial variation around this average (e.g. 2.5% of respondents reported their child to have no such books – 142 families). When it came to toddlers’ television watching, this was reported as very infrequent (30 min or less each day) by caregivers irrespective of whether this was solo or shared television watching. Again, there was notable variation to this. For example, 171 caregivers (3% of the baseline sample) reported four or more hours of shared television watching while 95 reported that their toddler watched 3 + hours of television by themselves every day.

Table 3 illustrates the bivariate associations between the eight measures of THLE and the two measures of toddlers’ television watching. The variation between the THLE and television watching items in terms of their averages is apparent within this table as are the topics (shared or otherwise) that the items focused upon. The responses to the items were most similar (via highest correlation coefficients) for items that shared a focus. For example, activities focused upon shapes and their names (items 3 and 4, rs = 0.49, p < 0.001), the frequency with which toddlers were read to and the number of books that they have (rs = 0.45, p < 0.0001), and how much time a toddler spends watching television either solo or shared with someone else (rs = 0.47, p < 0.0001). It is notable that the association between the television watching items and the THLE items, while negative, is very small (in terms of the magnitude of the correlation coefficients). This suggests that it is not the case that the toddlers in this sample experienced more frequent television watching in place of developmentally stimulating activities shared with adults, but rather that these are separate reported activities that all toddlers experience but on a more or less frequent basis.

Internal consistency

The internal consistency of four possible and alternative THLE scales was statistically appraised via Guttmann’s lambdas, Cronbach’s alpha, and McDonald’s omega (see Table 4). The four possible alternative THLE scales varied from one another as regards their inclusion of the two-toddler television watching items: included, alternatively excluded, and both excluded. Internal consistency was highest (across all seven statistical estimators of internal consistency) when the two television watching items were excluded - notably passing the commonly
accepted critical threshold value for $a$ of 0.70–and a finding that was in-keeping with the low correlations between the THLE and television watching items shown in Table 3. Notably, when each of the eight THLE items was considered for exclusion in order to increase the overall consistency of the other items, no increase in the values of the lambdas or omega coefficients were returned (which would indicate an opportunity for a scale comprised of items with greater internal consistency). This suggested that a THLE scale should be created that drew upon all eight of the THLE items and none of the items measuring toddler television watching.

The appropriateness of a unidimensional THLE scale comprised of the eight THLE items was then statistical evaluated through specification of a Confirmatory Factor Analysis (CFA) and appraisal of the fit of this model to the ECCE data to which it was applied. This CFA modified standard errors to take account of the nesting of families within SSCCs and, informed by the correlations above, specified correlated residuals between the two shapes and space items ($r = 0.23$, $p < 0.001$) and between the two items related to books ($r = 0.24$, $p < 0.001$; note that IRT modeling is not suitable when such correlations exist (DeMars, 2010) and that the modeling of correlated residuals in CFA is common practice though still subject to debate; see Bandalo, 2021). The resulting CFA demonstrated model fit that is commonly regarded as acceptable according to both the Root Mean Square Error Of Approximation (RMSEA) = 0.036 (90% confidence interval: 0.031 to 0.042) and the Comparative Fit Index (CFI) = 0.978 (e.g. Rigdon, 1996). Standardized factor loadings varied in magnitude from 0.33 (frequency of messy play) to 0.60 (frequency of being read to). The results of the CFA procedure confirmed the appropriateness of specifying a single THLE scale from the eight THLE items.

Once the items were identified that could consistently reflect a common response from caregivers as regards their toddler’s home learning environment, a THLE scale was created through summation of these (eight) items -- summation being used in order to create a THLE measure that matched the format of the PHLE measure upon which the THLE was based. With a THLE measure created, the final step undertaken in assessment of the reliability of the THLE was to appraise the THLE measure for its stratification by the age-range of the $n = 5,717$ children sampled by ECCE at baseline (when the average child was 14 months old). A small but significant association was found ($r = 0.20$, $p < 0.001$, $n = 5,687$) indicating a slight bias in the THLE scale: older children were somewhat more likely to receive a higher THLE score. In response, a child age-adjusted version of the THLE scale was developed by first regressing the THLE scale on the toddlers’ ages, saving the standardized residuals, and then $z$-scoring these residual values. The result was an age-adjusted THLE scale (adj.THLE) that while still highly correlated with the THLE scale ($r = 0.98$, $p < 0.001$, $n = 5,696$) was now zero correlated with the age of children who took part in the baseline ECCE survey of families ($r = 0.00$, $p = 1.000$, $n = 5,687$). Table 5 shows the means and distributions of the THLE, adj.THLE, and for comparison, the PHLE measure. For comparative purposes it is worth noting that the PHLE scale was not associated with the children’s ages, either at mean age 14 months ($r = 0.02$, $p = 0.262$, $n = 2,603$) or at mean age 38 months ($r = 0.02$, $p = 0.364$, $n = 2,604$).

### Measurement Validity

The association between children’s Home Learning Environments over time (appraising predictive criterion validity with convergent construct validity). With two versions of a THLE scale created (one via simple summation of the eight THLE items, the other adjusted for the sampled toddlers’ ages), the second stage of the statistical analyses considered the degree to which the THLE related to other measures in the ECCE dataset. First, the extent to which the THLE was associated with the PHLE was considered -- this to explore the consistency with which higher or lower THLE scores were likely to remain as such on average 24 months later (at mean child age 38 months). There was a moderate positive association shared between the THLE and PHLE over time ($r = 0.36$, $p < 0.001$, $n = 2,595$) with this equal in size and statistical significance to that shared between the age-adjusted THLE and PHLE ($r = 0.36$, $p < 0.001$, $n = 2,595$). Thus there was a tendency for a home learning environment to remain relatively stable–home learning environments around age 1 that

### Table 4

| Statistic | Including all ten items ($n = 5,688$) | Excluding watching TV on own ($n = 5,691$) | Excluding watching TV with adult ($n = 5,691$) | Excluding both watching TV items ($n = 5,696$) |
|-----------|-------------------------------------|------------------------------------------|---------------------------------|---------------------------------|
| $\lambda_1$ | 0.57                               | 0.59                                     | 0.59                           | 0.62                           |
| $\lambda_2$ | 0.66                               | 0.68                                     | 0.69                           | 0.72                           |
| $\lambda_3$, $\alpha$ | 0.63                              | 0.66                                     | 0.66                           | 0.71                           |
| $\lambda_4$ | 0.53                               | 0.59                                     | 0.59                           | 0.66                           |
| $\lambda_5$ | 0.65                               | 0.67                                     | 0.68                           | 0.71                           |
| $\lambda_6$ | 0.66                               | 0.67                                     | 0.67                           | 0.70                           |
| $\omega$    | 0.65                               | 0.68                                     | 0.68                           | 0.73                           |

### Table 5

| HLE measure | Mean | SD  | Skewness | Kurtosis | n    |
|-------------|------|-----|----------|----------|------|
| THLE        | 42.70| 7.68| -0.94    | 0.84     | 5,696|
| adj.THLE    | 0.00 | 1.00| -0.94    | 0.91     | 5,696|
| PHLE        | 30.59| 9.08| -0.21    | -0.49    | 2,604|
By looking at only the 2,608 families who participated in the longitudinal study, and not the 5,717 who took part in the baseline survey, a consistent sample is studied which makes for reliable comparison of demographic correlates of the THLE and PHLE.

**Tables 6, 7** show the results of the multilevel regression models that considered the extent to which demographic factors were associated with THLE scores. The results shown in **Table 7** broadly match those shown in **Table 6**, which was expected given the earlier correlation of \( r = 0.98 \) between the THLE and its child age-adjusted counterpart (see above). However, only **Table 6** (not **Table 7**) includes child age (at baseline assessment) as an independent variable—this as a statistical control to facilitate fairer comparison of demographic correlates across the two versions of the THLE scale. The demographic predictors of the THLE and child age-adjusted THLE scales are broadly equivalent. Toddlers were significantly likely to experience more stimulating home learning environments when their caregivers experienced less parenting stress, when their primary caregiver held higher academic qualifications, when families were not living in rent-

**TABLE 6** | Multilevel regression showing associations between demographics and the toddler home learning environment scale (THLE; \( n = 2,404 \) of the longitudinal \( n = 2,608 \)).

| DV | Independent variable | Model 0 | Model 1 |
|---|---|---|---|
| | B | S.E. | B | S.E. |
| THLE Intercept | 43.22*** | 0.30 | 27.54*** | 2.25 |
| Child gender: Male | 0.21 | 0.25 | 0.10 | 0.03 |
| Child’s age at “baseline” | 0.52*** | 0.05 | 0.10 | 0.45 |
| Mother’s age at “baseline” | 0.01 | 0.03 | 0.01 | 0.02 |
| (Concurrent) caregiver mental health/well-being | -0.01 | 0.02 | 0.04 | 0.04 |
| (Concurrent) parenting stress | -0.05*** | 0.01 | 0.26 | 0.33 |
| Highest caregiver qualification (equivalent): NVQ1 | 2.15 | 0.83 | 0.33 |
| NVQ2 | 3.12*** | 0.67 | 0.49 |
| NVQ3* | 0.05*** | 0.70 | 0.06 |
| NVQ4* | 4.05*** | 0.69 | 0.68 |
| NVQ5* | 4.95*** | 0.74 | 0.77 |
| Other* | 1.95* | 0.97 | 0.30 |
| Tenure: Rent it* | 2.77** | 0.98 | 0.43 |
| Shared ownership* | 2.79*** | 0.78 | 0.43 |
| Mortgage* | 3.09 | 1.28 | 0.37 |
| Own it outright* | 1.26 | 0.78 | 0.21 |
| Household income: \(< £4.999$$ | -0.66 | 0.89 | -0.10 |
| £5.000–£9.999$$ | -0.77 | 0.55 | -0.15 |
| £10.000–£19.999$$ | -1.34** | 0.44 | -0.21 |
| £20.000–£29.999$$ | -1.16** | 0.42 | -0.18 |
| £30.000–£39.999$$ | 0.09 | 0.43 | 0.01 |
| £40.000–£49.999$$ | 0.40 | 0.45 | 0.06 |
| Model fit: \(-2\loglikelihood | 17,438.16 | 15,615.14 |
| Model fit: \(\Delta-2\loglikelihood | 1823.02 |
| Intra class correlation | 0.15 | 0.10*** |
| Unexplained variance: | | |
| Sure start Children’s centers (SSCCs) | 8.10 | 4.29 |
| Families | 44.80 | 36.62 |
| Proportion of variance explained: | | |
| SSCCs | 0.47 |
| Families | 0.18 |

*p < 0.05; **p < 0.01; ***p < 0.001; DV, dependent variable; NVQ, national vocational qualification (equivalent level); †reference category, No qualifications; ‡reference category, Tenure (Rent free); ††reference category, household income > £50,000; intra class correlation, the proportion of variation in THLE scores attributable to differences between SSCCs rather than differences between families; †††Residual intra class correlation.

The statistical predictors of children’s Home Learning Environments (appraising criterion validity with discriminant construct validity). Although the range of caregiver responses to the THLE and PHLE were similar to one another and this provides evidence of the measurement validity of the THLE, further statistical comparison of the scales is possible to more thoroughly interrogate the measurement validity of the THLE. This was achieved by comparing the extent to which THLE and PHLE scores were statistically associated with a variety of demographic measures differentiating the 2,608 families who took part in the longitudinal survey of families (and following a similar approach to that of Toth et al., 2020 for older age groups within the EPPE study within which the PHLE was developed).
free accommodation, and where a family’s household income was higher. Further, THLE scores were consistently not associated with a child’s gender, a mother’s age, or the primary caregiver’s mental health/well-being. The largest association between THLE (either version) and family demographics was that shared with the primary caregiver’s highest held academic qualifications (by effect size). The second largest association was with child age at baseline assessment (Table 7)—hence providing further support for our consideration of a version of the THLE that was adjusted for the sampled toddler’s age in months.

Comparing the demographic correlates of THLE scores (Tables 6, 7) with PHLE scores (Table 8), both similarities and differences can be observed. Considering similarities first, the primary caregiver’s level of academic qualification remained the single largest statistical correlate of the extent to which children were likely to experience more stimulating HLEs. Furthermore, the more parenting stress that was experienced by a child’s caregivers (in either the toddler or preschool periods) the less stimulating the average HLE was likely to be. In addition, the extent to which an HLE was likely to be more or less stimulating was consistently not associated with either a mother’s age or the primary caregiver’s mental health/wellbeing. An over-arching similarity is that the proportions of variation in HLE scores that were explained by these predictors remained very similar across the age-adjusted toddler and preschool HLE measures. Around three times the variation in HLE scores that were explained by these predictors was variation between the families themselves (adj.THLE: 0.46 vs. 0.16; PHLE: 0.33 vs. 0.09). It is possible that this reflects the location of the sampled SSCCs in neighbourhoods classified as more disadvantaged but being open to all individual families who wished to use them (see Sylva et al., 2013).

Considering next the differences between the demographic correlates of the THLE and PHLE, while THLE scores were differentiated by child age, household income, and the tenure status of a family’s home, this was not found for the PHLE. Instead, PHLE scores were differentiated by child gender with the average boy likely to experience a significantly less stimulating preschool HLE than the average girl. The implications for the measurement validity of the THLE scale from these findings are mixed. The measurement validity of the THLE is supported by the similarities between the results, but the differences between the correlates of the THLE and PHLE scales require comparison.

| TABLE 7 | Multilevel regression showing associations between demographics and the child age-adjusted Toddler Home Learning Environment Scale (adj.THLE: n = 2,404 of the longitudinal n = 2,608). |
|---|---|
| DV | Independent variable | Model 0 | Model 1 |
| | | B | S.E. | B | S.E. | Effect size |
| Adj.THLE | Intercept | 0.06 | 0.04 | -0.11*** | 0.28 |
| | Child gender: Male | 0.03 | 0.03 | 0.04 |
| | Mother’s age at “baseline” | 0.00 | 0.00 | 0.00 |
| | (Concurrent) caregiver mental health/well-being | 0.00 | 0.00 | 0.00 |
| | (Concurrent) parenting stress | -0.01*** | 0.00 | -0.40 |
| | Highest caregiver qualification (equivalent): NVQ1* | 0.28* | 0.11 | 0.33 |
| | NVQ2* | 0.42*** | 0.09 | 0.49 |
| | NVQ3* | 0.54*** | 0.09 | 0.64 |
| | NVQ4* | 0.58*** | 0.09 | 0.68 |
| | NVQ5* | 0.67*** | 0.10 | 0.79 |
| | Other* | 0.27* | 0.13 | 0.32 |
| | Tenure: Rent it** | 0.37** | 0.13 | 0.44 |
| | Shared ownership* | 0.38*** | 0.10 | 0.45 |
| | Mortgage* | 0.31 | 0.17 | 0.37 |
| | Own it outright** | 0.19 | 0.10 | 0.22 |
| | Household income: < £4,999*** | -0.09 | 0.12 | -0.11 |
| | £5,000–£9,999*** | -0.11 | 0.07 | -0.13 |
| | £10,000–£19,999*** | -0.18** | 0.06 | -0.21 |
| | £20,000–£29,999*** | -0.15** | 0.06 | -0.18 |
| | £30,000–£39,999*** | 0.01 | 0.06 | 0.01 |
| | £40,000–£49,999*** | 0.05 | 0.06 | 0.06 |
| Model fit: Δ-2*loglikelihood | 6,838.90 | 5,903.45 |
| Intra class correlation | 0.15 | 0.10*** |
| Unexplained variance: | | |
| Sure start Children’s centers (SSCCs) | 0.13 | 0.07 |
| Families | 0.76 | 0.64 |
| Proportion of variance explained: | | |
| SSCCs | 0.46 |
| Families | 0.16 |

Note: *p < 0.05; **p < 0.01; ***p < 0.001; DV, dependent variable; NVQ, national vocational qualification (equivalent level); Highest PCG Q, highest primary caregiver qualification; *reference category, No qualifications; **reference category, tenure rent free; ***reference category, household income > £50,000; intra class correlation, the proportion of variation in adj.THLE scores attributable to differences between SSCCs rather than differences between families;Δ***residual intra class correlation.
to existing academic literature regarding age-related and gender variation in the known correlates of home environments. This is undertaken in the Discussion.

**DISCUSSION**

With the potential to serve as long-term predictors of a range of educational and developmental outcomes there is the requirement for early HLEs to be measured accurately. Further, while previous research has demonstrated the predictive power of the home learning environment in the preschool years (ages 3–5), less is known about the predictive power of the home learning environment for younger children. (Dodici et al., 2003). This paper responded to this issue by reporting the results from a series of analyses that investigated the reliability and validity of a new measure of the toddler home learning environment (THLE). This new measure was an adaption of a well-known existing measure of the preschool HLE developed by the longitudinal EPPSE investigation covering children age 3 + to 16 years in England (see Sylva et al., 2010).

Comprised of eight items that reflected the degree to which toddlers experienced a more or less stimulating home learning environment, the results presented in this paper show that higher THLE scores were more likely to be reported for older toddlers. However, the meaning of this association is informed by the content of the eight THLE items and by the rapid rate of developmental change in the toddler years, particularly as regards language development (e.g. Rodriguez and Tamis-LeMonda, 2011). The THLE was designed to be sensitive to adult-child interactions across the toddler period and this obliged the inclusion of items that better reflected stimulation for older toddlers (particularly relating to text and words). Thus, our finding of higher THLE score for older children is in-keeping with what one would expect: for older toddlers to more regularly experience activities related to texts and words as part of appropriate developmental scaffolding (see Granott, 2005).

However, there is a practical consequence from the association between THLE scores and the age of toddlers for how the THLE scale should be used in future. In circumstances where there is a wide age-range of toddlers whose home learning environments are to be assessed, the tendency for the THLE scale to return higher scores for older toddlers should be taken into account. This paper demonstrated two ways for this to be carried out that would lend themselves to future research: modification of the

---

**TABLE 8 | Multilevel Regression showing demographic correlates with the preschool home learning environment scale (PHLE; n = 2,409 of the longitudinal n = 2,608).**

| DV                  | Independent variable | Model 0 |        | Model 1 |        | Effect size |
|---------------------|----------------------|---------|--------|---------|--------|-------------|
|                     |                      | B       | S.E.   | B       | S.E.   |             |
| PHLE                | Intercept            | 30.49***| 0.26   | 31.91***| 0.34   | -0.26       |
|                     | Child gender: Female | -2.23***| 0.35   | -0.35   | 0.02   |             |
|                     | Child’s age at “outcome” | 0.03   | 0.06   | 0.03   | 0.02   |             |
|                     | Mother’s age at “baseline” | -0.01 | 0.03   | -0.03   | 0.01   |             |
|                     | (Concurrent) caregiver mental health\well-being | -0.02  | 0.08   | -0.08   | 0.01   |             |
|                     | (Concurrent) parenting stress | -0.11***| 0.03 | -0.20   | 0.01   |             |
|                     | ‘Highest caregiver qualification: NVQ1’ | 0.44   | 1.13   | 0.06   | 0.05   |             |
|                     | NVQ2’                | 1.02    | 0.92   | 0.12   |        |             |
|                     | NVQ3’                | 3.36***| 0.94   | 0.40   |        |             |
|                     | NVQ4’                | 4.68***| 0.92   | 0.55   |        |             |
|                     | NVQ5’                | 5.26***| 0.99   | 0.62   |        |             |
|                     | Other’               | 3.52**  | 0.42   |        |        |             |
|                     | Tenure: Rent it”     | 1.51    | 1.37   | 0.18   |        |             |
|                     | Shared ownership”    | 1.10    | 1.12   | 0.13   |        |             |
|                     | Mortgage”            | 0.50    | 1.86   | 0.06   |        |             |
|                     | ‘Own it outright”    | 0.58    | 1.12   | 0.07   |        |             |
|                     | Household income: < £4.999” | -0.10 | 1.24   | -0.01  |        |             |
|                     | £5.000–£9.999”       | 0.11    | 0.75   | 0.01   |        |             |
|                     | £10.000–£19.999”     | 0.38    | 0.60   | 0.04   |        |             |
|                     | £20.000–£29.999”     | -0.03   | 0.57   | 0.00   |        |             |
|                     | £30.000–£39.999”     | 0.37    | 0.59   | 0.04   |        |             |
|                     | £40.000–£49.999”     | 0.85    | 0.62   | 0.10   |        |             |

Model fit: −2*loglikelihood | 18,845.21 | 17,192.67 | 0.04*** |
Model fit: Δ −2*loglikelihood | 1,652.54 | 1.04*** |
Intra class correlation | 0.05 | 0.04*** |
Unexplained variance: | 4.20 | 2.81 |
Sure start Children’s centers (SSCCs) | 78.67 | 71.61 |
Families | 0.33 | 0.09 |

*p < 0.05; **p < 0.01; ***p < 0.001; DV, Dependent Variable; NVQ, National Vocational Qualification (equivalent level); reference category, No qualifications; reference category, Tenure (Rent free); reference category, household income > £ 50,000; intra class correlation, the proportion of variation in PHLE scores attributable to differences between SSCCs rather than differences between families; **residual intra class correlation.
THLE scale or inclusion of child age in any multivariate statistical analyses that are subsequently undertaken. A third (simpler) alternative (potentially for use outside of research projects) would be for interpretations of THLE scores to demonstrate explicit understanding of the fact that THLE scores are likely to be higher for older toddlers. Of course, the inverse is also true: should the THLE be used with a narrow age range of toddlers then there is less of a need to take age-related variation in THLE scores into account in either THLE measurement, use in statistical analyses, or in subsequent interpretation. That said, less need does not imply no need.

The need for the measurement validity of the THLE to be appraised through a combination of statistical results and existing academic literature also extends beyond the association that the THLE shared with the age of toddlers --particularly given the academic literature also extends beyond the association that the THLE shared with the age of toddlers and the PHLE. On one hand, the results of this paper showed a moderate degree of consistency between the THLE and PHLE. For example, a child’s THLE score was shown to be broadly in line with their PHLE score twenty four months later, parenting stress and PCG academic qualifications were consistently associated with HLE scores at both ages, and PCG mental health/well-being was consistently not associated with either the THLE or PHLE. Such consistency (with the prior validated THLE scale or inclusion of child age in any multivariate statistical analyses) was consistently found to relate to scores on the (prior validated) PHLE - especially given that the THLE was an adaption of the PHLE.

Considering limitations, first the data comes from parental self-report and thus is subject to the biases inherent in this form of data (e.g. social desirability in the responses given—not uncommon in measures of the HLE; e.g. Lee and Xie, 2017; Peacock-Chambers et al., 2017). Second, the data lacked long-term educational and developmental outcomes by which to appraise the THLE relative to the PHLE. The project from which the data came (and within which the THLE was developed) was not designed to last long enough to obtain this data–certainly not the twelve years of the prospective longitudinal EPPSE investigation within which the PHLE was developed. Third, the Cronbach alpha estimate was at the lower-end of those commonly treated as acceptable. Fourth, no information was collected about tablet or smart phone use (including apps) by toddlers or pre-school children. This reflects the historic context when the ECCE study took place (such technology being much less common in 2010 when the first children were recruited to the evaluation). Although ECCE investigated television viewing alongside the HLE (as a kind of ‘displacement activity’; Dore, et al., 2020), future research could also explore the extent to which tablet or phone use is related to other kinds of home learning activities for this very young age group. This would also help to inform the currently sparse and at-times contradictory evidence concerning how these devices (and apps on them) can benefit and/or hinder child development (e.g. Hall et al., 2019a).

Implications for Researchers, Early Years Practitioners, and Policy Makers

Given that this paper provides new evidence on the reliability and validity of the THLE scale, there are a number of implications for researchers, early years practitioners, and early years policy makers. First, there is much scope for further exploration of the long-term correlates of the THLE measure in future academic research (beyond preschool-period: HLE, verbal cognitive abilities, non-verbal cognitive abilities, and prosocial behaviors). One of the strengths of the PHLE measure (upon which the THLE is based) is that it has shown itself to be predictor of developmental and educational outcomes above and beyond socioeconomic status (e.g. Sammons et al., 2014; Sammons et al., 2015b) – a rare finding in studies of child development and educational progress. Ideally, the THLE would also be able to show such long-term predictive power—even if indirectly through
first predicting the PHLE. However, for this to be investigated the THLE needs to be included in a future prospective longitudinal study, ideally alongside the PHLE. The inclusion of the eight THLE items in future research could be valuable given that the PHLE has already been included in many post-EPPSE studies (e.g. the Millennium Cohort Study [MCS] and the Study of Early Education and Development [SEED], and the BiKS study in Germany).

Second, future research into the effects of home learning environments on children’s development and learning would benefit from a critical reflection upon the common practice of creating HLE scales via summation or averaging of HLE items, particularly as these approaches oblige HLE items to make equal contributions to an HLE scale. Alternative approaches to combining items into a scale avoid this forced equal contribution from items and this can result in a scale with greater predictive power (e.g. Hall et al., 2010). There are a wide range of statistical factor analysis techniques that can be used for this purpose (examples of which can be found used in HLE papers; e.g. Linberg et al., 2020) and these techniques are continually developing (e.g. ‘Exploratory Structural Equation Modeling’; see Marsh et al., 2010). Thus, a systematic examination of the impacts of using these alternative approaches in the construction of HLE scales has the potential to improve their reliability and validity. In turn, improved HLE scales have the potential to offer fresh insights into child development and the predictors of child development.

Third, the THLE --although subject to limitations -- shows merit as tool for use in developing a fuller understanding of the environments within which toddlers learn, grow, and live. As such, the THLE can serve as a useful tool for those working with vulnerable families--particularly for researchers and practitioners whose work is concerned with children’s development and with adult-child interactions. We also believe that the THLE will be particularly useful for those who work within or alongside early interventions such as Sure Start Children’s Centers and its international equivalents (e.g. Head Start in the United States of America; Dream Start in South Korea, and Family Centers in Germany; see respectively: Welshman, 2010; Lee et al., 2015; Stöbe-Blossey et al., 2009). However, the use of the THLE outside of the cultural context in which it was developed (England) prompts the need for more research to investigate the potential need for localization and adaption.

Fourth, for policy makers in the early years, the extension of the well-known PHLE scale down to the toddler years provides new opportunities to evaluate early years practice (across education, health and social policy) against criteria of demonstrable efficacy, effectiveness, and cost-effectiveness (e.g. Heckman, 2006). For example, to show that Sure Start Children’s Centers (as a social policy with variable localized implementation) have the potential to significantly improve the home learning environments of the toddlers whose families use them (see Sammons et al., 2015a; Hall et al., 2019b).

**CONCLUSION**

This paper moves the HLE literature forward with the provision of evidence toward the reliability and validity of a new parental self-report scale of the HLE supporting the development of toddlers. The new THLE scale has the potential to inform research, practice and policy (in the toddler period, pre-schooler period, and later periods) by prompting an increased understanding of the early drivers of educational attainment and development. However, more research is required, particularly on what (if any) long-term outcomes can be expected from home learning environments in the toddler period, and what the implications are for home learning environments from different forms of smart phone and tablet use with children at this very young age.

**DATA AVAILABILITY STATEMENT**

The data analyzed in this study is subject to the following licenses/ restrictions: The ECCE dataset is owned by the United Kingdom Government. Requests to access these datasets should be directed to J.E.Hall@Soton.ac.uk.

**ETHICS STATEMENT**

The studies involving human participants were reviewed and approved by the University of Oxford Central University Research Ethics Committee (CUREC). Written informed consent to participate in this study was provided by the participants’ legal guardian/next of kin.

**AUTHOR CONTRIBUTIONS**

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

**FUNDING**

This work was supported by a grant from the United Kingdom Department of Education (grant number EOR/SBU/2008/023).

**ACKNOWLEDGMENTS**

Our thanks to Sina Fackler (sina.fackler@lifbi.de) for early analyses in preparation for this paper.

**REFERENCES**

Abidin, R. R. (1995). *Parenting Stress Index: Professional Manual.* 3rd ed. Odessa, FL: Psychological Assessment Resources, Inc.

Anders, Y., Grosse, C., Rossbach, H.-G., Ebert, S., and Weinert, S. (2013). Preschool and Primary School Influences on the Development of Children’s Early Numeracy Skills between the Ages of 3 and 7 Years in Germany. *Sch. Effectiveness Sch. Improvement* 24 (2), 195–211. doi:10.1080/09243453.2012.749794
