



Education
Endowment
Foundation

Update of the EEF's Early Years Measures Database

Protocol

October 2025

Sara Bashir Malik, Lydia Marshall, Anushay
Mazhar, Astrid Pickenpack, Rachel Outhred and
Robert Wishart





Education
Endowment
Foundation

The Education Endowment Foundation is an independent charity dedicated to breaking the link between family income and education achievement. We support schools, colleges and early years settings to improve teaching and learning for 2 – 19-year-olds through better use of evidence.

We do this by:

- **Summarising evidence.** Reviewing the best available evidence on teaching and learning and presenting in an accessible way.
- **Finding new evidence.** Funding independent evaluations of programmes and approaches that aim to raise the attainment of children and young people from socio-economically disadvantaged backgrounds. Putting evidence to use.
- **Putting evidence to use.** Supporting education practitioners, as well as policymakers and other organisations, to use evidence in ways that improve teaching and learning.

We were set-up in 2011 by the Sutton Trust partnership with Impetus with a founding £125m grant from the Department for Education. In 2022, we were reendowed with an additional £137m from government, allowing us to continue our work until at least 2032.

For more information about the EEF or this report please contact:



Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP



0207 802 1653



info@eefoundation.org.uk



www.educationendowmentfoundation.org.uk



Department
for Education

What
Works
Network

Table of contents

Summary	3
Background and context	4
Project Aims and Objectives	5
Methodology	5
Refining the EYMD taxonomy and structure	6
Updating the EYMD content and developing guidance and resources to capture learning	10
Revising the EYMD design interface	16
Data protection	17
Risk management	17
Team	20
Conflicts of interest	20
Timeline	20
References	22
Annex A: Draft taxonomy	23
Annex B: Search terms	31

Summary

The Education Endowment Foundation (EEF) has commissioned Oxford MeasurEd to update the Early Years Measures Database (EYMD), first launched in 2017. The original EYMD provided evaluators and researchers with a structured overview of tools for assessing children aged 0-6 across developmental domains such as language, social-emotional development, literacy, and numeracy. However, shifts in policy, research priorities, and practice over the past decade, including revisions to the Early Years Foundation Stage (EYFS) framework, increased demand for early years evaluations, and new government priorities around assessment of very young children, have highlighted the need for a comprehensive update. A particular priority is to strengthen coverage of the 0-3 age group, which was underrepresented in the original EYMD despite being a period of rapid brain development and heightened sensitivity to environmental influences. Valid and reliable measures for this age range remain relatively scarce, yet they are critical for evaluating interventions and informing practice in line with EEF's expanded remit and wider government focus on early development.

The update aims to:

- Expand and refine the EYMD taxonomy of domains and sub-domains to ensure closer alignment with the EYFS framework.
- Identify and appraise both new and existing outcome measures, including those suitable for children aged 0-3, with systematic attention to their psychometric properties, administration requirements, and appropriateness for use in UK evaluations.
- Improve the usability and presentation of the EYMD, including revisions to the database design interface and the development of a more intuitive and constructive rating system for psychometric robustness and implementation feasibility.
- Provide guidance and practical resources to ensure the EYMD can be used effectively by evaluators, researchers, and other stakeholders.

The methodological approach integrates three strands of work: refining the taxonomy and structure of the database, updating the content and related information on the EYMD, and revising the EYMD design interface. The project will incorporate stakeholder consultations, evaluator workshops, and user testing to ensure that revisions reflect both expert knowledge and practitioner needs.

The updated EYMD will provide a more comprehensive and user-friendly resource and strengthen the EEF's role in supporting high-quality early years research and evaluation. The primary audience for the EYMD are the EEF and evaluators selecting outcome measures for use in EEF-funded trials in England. Secondary audiences include other researchers and evaluators measuring early years outcomes in the UK and beyond, as well as practitioners and funders seeking guidance on robust outcome measurement practice.

This protocol sets out how the project will be conducted to achieve these objectives. The protocol is intended as a record of the work to be carried out by Oxford MeasurEd to update the EYMD in 2025, and does not capture future updates to the EYMD that might be carried out by the EEF or others. However, it does explain the guidance Oxford MeasurEd will provide on how the EYMD can be updated.

Background and context

Following the Education Endowment Foundation (EEF)'s decision to expand its remit of work to the early years sector in 2014, the EEF commissioned the Institute of Education at University College London (UCL) to carry out a review of measures designed to assess 0-6-year-old children (Dockrell et al. 2017). This resulted in the launch of two key resources in 2017 – (1) the **Early Years Measures Database** (EYMD), which catalogues tools for assessing developmental domains including language and communication, social and emotional development, literacy and numeracy, and (2) an **accompanying document**, which summarises the findings of the review and how the EYMD was developed, and offers some guidance on what to consider when choosing measures to evaluate children's abilities. It also discusses some measures for the home and early learning environments, which were not included in the EYMD, and includes information on the number of measures for each domain and the range of competencies they assess. This review was an initial attempt to identify measures that would be most relevant to EEF trials in the early years and was not intended to be comprehensive in its coverage of how outcomes for early years should be measured.

A decade after commissioning the initial review of measures for the early years, the EEF carried out an internal review to explore whether there was a need to update the original EYMD and what this update would look like. Feedback from different stakeholders in the internal review highlighted gaps in coverage, usability challenges and the need to expand its scope, showing that the database in its current form did not meet research needs or support practice effectively. These findings underlined the need for an update to ensure the EYMD remains a relevant, effective and a comprehensive resource. At the same time, it was widely acknowledged that the EEF had the opportunity to elevate the database into a definitive resource for evidence-based assessments, establishing EEF as a leader in the early years research community.

In 2022, the EEF received a re-endowment by the DfE which extended its early years remit both in terms of the age range of children its work could support (2 year olds+) and the type of early years settings it aimed to work with (school-based, PVI and childminder settings). As the original EYMD had limited coverage for two-year-olds and was primarily focused on school-based contexts, this expansion created a clear need for additional measures that could capture outcomes for younger children and across a wider range of early years settings. This extension, along with a collaboration with Department for Education's (DfE) Stronger Practice Hubs to support education recovery following Covid-19 pandemic, notably increased the number of early years evaluations commissioned by EEF¹. This work was the start of a shift in the organisation's strategic priorities to focus more strongly on the early years. EEF's early years strategy includes a focus on building expertise across the EEF's evaluator panel in conducting early years evaluations, and aims to facilitate high-quality research through the development of better evaluation infrastructure. The development of the **Early Years Evidence Store** in 2022 and the **Early Years Toolkit** further underline the EEF's commitment to producing high-quality evidence for researchers and practitioners in the early years over the past decade.

Government's role and emphasis on the early years has also shifted in recent years. Since the release of the EYMD in 2017, there have been several updates to the DfE's **Early Years Foundation Stage (EYFS) framework** (DfE 2021), with an updated framework coming into effect in September 2025 (DfE, 2025a). The EYFS framework sets out the statutory requirements for teaching and learning in the early years, along with broader requirements for provision, and is updated annually. Within this, at present, the EYFS Profile, which is the statutory assessment at the end of the Reception year, includes 17 Early Learning Goals (ELGs) which outline the expected level of development that most children will achieve by the end of their Reception year.

Alongside these changes to the EYFS, wider government investment and policy initiatives have further underlined the centrality of early years. The 'Best Start in Life' strategy has placed emphasis on improving early development and ensuring more children reach a good level of development by age five in line with the government's Opportunity Mission which aims to increase the proportion of children achieving this to 75% by 2028. Starting in September 2025, around 80% of early education and childcare hours will be government-funded rather than parent-funded (IFS, 2024), marking a major expansion of state investment in the sector. The current Government's focus on the early years is also evident in the DfE's new areas of research interests (DfE 2025b), which call for more evidence on how best to develop tools to assess children's development before they go to school.

¹ At the time of writing, the EEF has completed 34 evaluations in the early years space, with a further 18 ongoing.

Alongside these policy shifts, a growing body of evidence underscores the importance of the earliest years of life. Development in the first three years of life is widely recognised as critical for later outcomes (see for example Tierney and Nelson III, 2009). Early development is driven by rapid neural connections shaped by experience, environment, nutrition, safety, and responsive stimulation. These foundations influence lifelong learning and wellbeing (*ibid.*), as well as other outcomes including earnings in adulthood (Britto et al, 2017). Yet despite this, valid and reliable measures for the 0-3 age group remain relatively scarce, leaving a major gap in the EYMD.

Taken together, these policy developments, funding reforms, and scientific insights strengthen the case for updating the EYMD. In particular, they point to the need both to expand coverage for the youngest children, an age group underrepresented in the original database, and to ensure alignment with the EYFS framework so that the EYMD remains relevant for policymakers and practitioners.

In light of this context and these evolving priorities, the EEF has commissioned Oxford MeasurEd to update the EYMD.

Project Aims and Objectives

The overall aim of the project is to update the original EYMD. Specifically, the project has four objectives, including:

1. **Defining domains:** The update will expand the EYMD to include and clearly define additional domains and refine existing domain labels, with the aim to ensure better alignment with other EEF EY resources such as the Early Years Evidence Store and national frameworks including the EYFS framework.
2. **Identification and review of new and existing measures:** The project will identify a new longlist of measures to add to the database, including suitable measures to assess children aged 0-3. For a shortlist of measures for which the information is available, the update will include an assessment of their psychometric properties, administration characteristics, use of tool, alignment with **EYFS Early Learning Goals**, and presence in any national administrative database. The project will also update the information on all existing measures in the database to align with the new structure of the database and ensure that the information is up to date. The review will examine how well different developmental domains and age groups are covered by existing measures, identifying gaps where measures are lacking. The suitability of internationally developed measures will be explored where UK-specific measures are lacking.
3. **Improving usability and presentation of EYMD:** The project will involve making recommendations to improve the usability of the EYMD, including the database interface, individual measures' pages, and rating systems for psychometric properties.
4. **Developing guidance and capturing learning:** The update will be accompanied by resources and mechanisms to guide EEF and its partners in effectively using the EYMD, including practical tips and insights from previously published EEF-funded trials. The project will also produce guidance for EEF on how to capture such insights moving forward.

Methodology

Our methodological approach encompasses three strands of work:

1. Refining the EYMD taxonomy and structure (Project aim 1)
2. Updating the EYMD content and developing guidance and resources to capture learning (Project aims 2 and 4)
3. Revising the EYMD design interface (Project aim 3)

The three work strands are described in the following section. They were informed by a review of key documents, including:

- EEF's initial review of EY assessment measures carried out by UCL (Dockrell et al. 2017)
- EEF's internal review of the EYMD in 2024
- EEF's Early Years Evidence Store (EYES) and accompanying resources

- the EYFS framework and Early Learning Goals

Refining the EYMD taxonomy and structure

Aims

The aims of this work strand support overall project objective 1 (see page 5). The aims are twofold: first, to align the EYMD national EYFS framework and other EEF early years resources such as the EYES; and second to refine the taxonomy underpinning the EYMD to include and clearly define new domains and refine existing domain definitions.

Approach

Alignment of outcome domains with the EYFS framework and other EEF resources

At inception, we agreed with the EEF that alignment with the EYFS framework should be a priority. In line with this, the EYFS provided the primary structure for the database. While this approach prioritises alignment with national policy frameworks above other theoretical frameworks, later stages of development including consulting evidence from recent literature and global resources, as well as with stakeholders, to refine definitions and ensure the taxonomy underpinning the framework was theoretically sound (see below). The domains outlined in Figure 1 show how the original domains in the EYMD map on to our revised domains that are aligned with the EYFS.

Figure 1 Identifying domains of interest²

The original development of the EYMD was based on detailed definitions of the learning domains rooted in key developmental literature (see Dockrell et al., 2017). We therefore aimed to build on these strong foundations. We cross-referenced the original descriptions with the EYFS statutory framework and the EEF's EYES to come up with initial definitions for the new domains, before turning to the wider literature (see below).

Targeted desk review

² In the 2017 EYMD review, measures that “tapped skills related to at least two of the covered domains” (Dockrell et al, 2017, p. 33) were labelled under a separate “omnibus” label. Measures previously labelled as “omnibus” will be retained in the updated database, but linked to specific learning/development outcome domains, with an indication that they measure more than one domain. This is because many established early years assessments span multiple domains, and the label omnibus will not help evaluators to find a measure relevant to the particular outcomes they are interested in. This approach will allow for more targeted filtering and ensuring better alignment with users’ priorities, such as selecting measures that best match the specific developmental outcomes they aim to evaluate or improve.

We carried out a targeted desk review to further refine our domain definitions and ensure they were comprehensive and reflected EEF's priorities and the evidence base on early child development. To do this, we consulted recent evidence reviews commissioned by the Education Endowment Foundation (EEF), including Early Years and Key Stage 1 Mathematics Teaching: Evidence Review (Hodgen et al., 2020) and Early Language Development: Needs, Provision, and Intervention for Preschool Children from Socio-Economically Disadvantaged Backgrounds (Law et al., 2017). We also drew on global frameworks, notably the World Bank's Toolkit for Measuring Early Childhood Development in Low- and Middle-Income Countries (Fernald et al., 2017) alongside other peer-reviewed literature relevant to early years outcomes and outcome measurement. The updated definitions can be found in Annex A.

This desk review also informed our development of a more detailed taxonomy to underpin the EYMD. Based on the original EYMD, the EYFS statutory framework, the EEF's EYES and the other literature reviewed, we developed a more detailed structure for each domain, comprising a set of sub-domains and associated skills. Sub-domains refer to distinct, theoretically grounded components within a broader developmental area, while skills represent observable behaviours or capacities within each sub-domain that might be measured through assessment tools. The lists of skills are not intended to be exhaustive, but aim to capture key skills that interventions may seek to develop and that evaluators may be seeking to measure. This structure aims to enhance clarity, support consistent categorisation of assessments, and help users better understand what is being measured and how it maps onto children's developmental trajectories.

Stakeholder consultation on taxonomy

To validate our definitions and the proposed domain structure, we conducted a two-stage consultation process with early years experts. These included both subject/domain specialists and general early years experts identified by EEF. First, we distributed an online survey to a 26 researchers to gather initial feedback on the proposed domains, sub-domains, and associated skills. Respondents were asked to comment on the clarity, relevance, and completeness of each domain, as well as to suggest any missing or overlapping areas. We received 11 responses to the survey.

We made changes to the draft taxonomy based on the survey feedback. Then, we held four consultation workshops, inviting the same pool of experts to attend. These workshops provided an opportunity to discuss the revisions we had made based on the survey as well as any areas of divergence in the feedback. We used feedback from the workshops to revise the domain definitions and taxonomy, ensuring that they reflected current evidence and aligned with current policy, practice and developmental theory.

Due to the timeline for the project, the workshops needed to be conducted during school holidays. This was a limiting factor, and a total of three experts were able to take part in the consultation workshops. During the process of finalising the draft taxonomy, we conducted the following additional consultations:

- Discussions with theme leads at the EEF to further refine the sub-domains, skills and definitions under the Early Mathematics and Personal, Social and Emotional Development domains. These theme leads have expertise in the particular subject areas. Consulting with them helped to ensure that the taxonomy reflected practice expertise as well as priorities for the EEF's work in these areas.
- A consultation with an expert in very early learning and development, to ensure that the taxonomy is relevant and applicable to very young children (aged 0-2).

Agreeing decisions with EEF

We presented the proposed domain structure, including revisions based on expert consultation, to the EEF for review. We also discussed areas of diverging feedback, which tended to be around the level of granularity to be included in the taxonomy, and how to represent where skills in one domain laid the foundations for skills in another domain, rather than any disagreement on which skills and outcome domains are important in early childhood. Following feedback and discussion, we agreed the structure and confirmed it as the framework to guide the next stage of the review. In the process of developing this protocol, we also received and addressed feedback from early years experts at EEF on the structure of and definitions in the taxonomy.

The following decisions were agreed upon:

- A more detailed structure will be used for the taxonomy underpinning the EYMD, with each domain comprising a set of clearly defined sub-domains and associated skills (see explanation above).

- Using this taxonomy, all measures in the database will be tagged with the relevant domain(s), sub-domain(s) and skill(s). Users will be able to filter and search the database by these categories, as well as other key factors such as age of the target population of children.

A draft of the full taxonomy of domains, sub-domains and skills was agreed and is set out in **Error! Reference source not found.** overleaf and in Annex A.

Further development

We note that the taxonomy will be iterated and finalised throughout the duration of the project. We may update the key skills lists as we review which skills are commonly measured within each sub-domain. As mentioned above, we do not intend for these lists of skills to be comprehensive, but aim for them to represent a) the key skills emphasised in the EYFS, our targeted desk review and consultations with EEF and other stakeholders as being important for early education provision and interventions; and b) the key skills that available measures target. In particular, the Understanding the World, Expressive Arts and Design and Digital Technology domains are less commonly discussed in the literature, and may well be refined as we review available measures.³ We will outline any changes made to the taxonomy following the publication of this protocol, in the final report presenting findings from the review.

Deliverables

The outputs of this strand are the taxonomy presented overleaf and this Protocol, describing how the database will be updated. Our intention is that the taxonomy provides a consistent framework to help users to select tools that best align with intended outcomes, compare measures and identify gaps in coverage.

³ We will not remove skills from the taxonomy because there are no measures available. Rather, we may add skills that are commonly measured within the target domains, or may adjust the naming of skills and domains to reflect how they are described in measurement documentation.

Figure 2: Draft taxonomy

The largest shapes represent domains, with sub-domains emanating from them and skills linked to the sub-domains. Some skills are linked to more than one sub-domain, shown with lines going from the skills to more than one sub-domain. Where a skill is attributed to more than one domain, we have used graduated colouring to illustrate this. Digital Technology skills and the Early Learning Environment are presented as underpinning learning and development in other domains.



Updating the EYMD content and developing guidance and resources to capture learning

Aims

The aims of this work strand support overall project aims 2 and 4. The aims are threefold:

- identifying a long list of new measures to add to the database, including suitable measures to assess children aged 0-3
- systematically documenting information for shortlisted measures, those with available psychometric properties, administration guidelines and scoring criteria, covering appropriate uses, alignment with EYFS Early Learning Goals, and presence in any national administrative database (for new measures and those already included in the existing EYMD)
- mapping measurement coverage and exploring the suitability of internationally developed measures where UK-specific measures are lacking

Approach

Sourcing a longlist of outcome measures

The first step will focus on developing a longlist of publicly available outcome measures in the domains of interest. The purpose is to identify:

- measures relevant to existing domains that have been developed since the 2017 review
- measures available for each new domain
- measures relevant to the new definition where existing domains have been expanded

We will use broad inclusion criteria to ensure comprehensive coverage of outcome measures while maintaining relevance. We have agreed these criteria with the EEF to ensure that resources are being spent on identifying measures that will be most useful to the end users. The criteria being used for inclusion at this stage include that the measure:

- targets one or more of the **domains** in the updated EYMD taxonomy
- published since the **2017** review
- suitable for children aged **0-6-year-olds**
- available in **English**

We have devised a search strategy to efficiently identify the most relevant measures while achieving breadth in the list of measures we identify. We will draw on a range of sources to maximise coverage, including in:

- outcome measures used in **EEF- and DfE funded early years evaluations**
- assessments identified during the **initial desk-based review** while restructuring the database (see *Refining the EYMD taxonomy and structure* above)
- **targeted manual searches** of established assessment publishers (such as Pearson and GL Assessment)
- the **World Bank Early Childhood Development Measurement Inventory**
- **recommendations** from EEF, the early years experts we have consulted with, and our internal Advisors

- relevant academic research databases (PsycINFO, SCOPUS and ERIC) using the search terms set out in Annex B^{4,5}

Throughout this process, we will pay particular attention to identifying appropriate measures for children aged 0-3, reflecting the EEF and DfE's growing interest and gaps in provision for this age group.

The longlist stage is not intended to provide in-depth evaluation of the measures but rather to capture a concise set of essential information for each measure. Therefore, the longlist will only capture essential information about each measure, including:

- assessment name, acronym and version
- summary of the test obtained directly from the author/publisher, if available
- domain(s) and sub-domain(s) targeted by the assessment
- age group
- publisher/author name and link to assessment

This information will be stored in the database about all assessments identified in the longlist, but further information and appraisal will only be provided for those that will be shortlisted based on pre-specified criteria (see *Screening* below). The summary of the longlisted measures from the author/publisher will be recorded verbatim and the specific skills assessed by the measure will not be identified.

Screening the longlist to produce a shortlist

We will screen the longlist of measures to produce a shortlist using criteria informed by the 2017 review, which provided a strong foundation for identifying robust, psychometrically sound, and practically feasible tools. However, rather than applying the 2017 requirement for UK-based standardisation samples as a strict initial filter, we agreed with EEF to consider this at a later stage in the appraisal process (see Table 1). This approach allows us to retain internationally developed tools that have documented use in UK settings, particularly valuable in domains such as learning environments, where some of the most well-established early years measures originate from outside the UK.

Table 1: Shortlisting criteria 2017 and 2025

Criteria	Original requirement	Updated requirement
UK standardisation sample	Required for norm-referenced assessments	Relaxed – documented use in UK settings accepted even if not UK-normed
Validity information	Required	Required
Reliability information	Required	Required
Administration guidelines	Required	Required
Scoring criteria	Required	Required

If more measures are available than we have resources to appraise, we will prioritise measures for inclusion in the database in consultation with EEF. We will work with EEF to agree prioritisation criteria, applying additional screening where needed. For instance, it might make sense to refine the inclusion criteria to shortlist those measures that are likely

⁴ If there are more measures available than we have resource to appraise, we will agree with the EEF an approach to prioritising the resources available within the project budget. This could include using systematic review software (EPPI Reviewer) or capping the search at a number of search results pages that allows to capture the most prominent and widely used measures. The approach used will be acknowledged in our methodology, as it may lead to the exclusion of some relevant tools. Alternatively, if we have more resources available after we have completed the appraisal process, and/or there are limited measures available for a certain domain or sub-domain, we will expand the scope of our search to find additional measures by using forward- and backward-citation tracking the literature identified in the search and/or to include measures published prior to 2017 that have not already been included in the review.

⁵ We will outline any changes made to the search strings following the publication of this protocol, in the final report presenting findings from the review.

to be most relevant and most feasible for use in EEF's evaluations. We may tighten filter criteria for inclusion for some domains or sub-domains where there are more measures available, and not for others. Any adjustments to the inclusion and search criteria will be documented in an update to this protocol.

If measures from the long list do not pass screening, the reason(s) for not being shortlisted will be recorded in the database. The reasons for exclusion would include:

- validity information not available
- reliability information not available
- administration guidelines not available
- scoring criteria not available

Appraising the psychometric properties of shortlisted outcome measures

We will review the properties of the shortlisted measures using adapted criteria based on the European Federation of Psychologists' Associations (EFPA) review model Version 4.2.6 (EFPA, 2013). The EFPA review model offers a detailed structure for evaluating educational and psychological assessments. It covers both psychometric properties and qualitative implementation considerations.

Table 2 sets out the rating system we will use to rate key psychometric properties of each measure. We will rate the evidence of the robustness of the property as limited (✓), moderate (✓✓) or strong (✓✓✓). Further information will also be documented against each of these criteria to contextualise and justify the rating (e.g. the reliability test and coefficient, as well as the rating). These criteria have been drawn from the EFPA review model. However, whereas the EFPA applies a four-point scale for most criteria, we have adapted this into a three-point rating system. The criteria for awarding a 'strong' rating were drawn directly from the EFPA review model. The 'limited' and 'moderate' criteria were then developed by progressively relaxing these standards, informed by our team's assessment expertise and professional judgment.

We will also document any other evidence about standardisation and further aspects of validity and reliability. This will include the sample size and characteristics of any standardisation exercises outside of the UK where a measure has not been standardised in the UK, as well as the response rates in standardisation exercises – a key factor affecting representativeness. Turning to validity, we recognise that a test's validity may vary depending on the intended use, the population tested, and the type of decision being made. We will aim to capture and report where specific validity evidence exists for different uses or populations (e.g., screening vs. diagnostic use, or general vs. SEND populations). Where evidence is uneven or limited to particular subgroups, we will note this in our narrative synthesis and flag it in the database where possible. We will also record any evidence of issues with floor and ceiling effects, and skewness.

To evaluate the properties of each measure, we will review information provided in publisher manuals, official websites, and peer-reviewed academic publications. This includes standard administration and technical manuals intended for users, which may contain reliability, validity, and standardisation data. Publishers may hold additional data not publicly accessible, but our review will rely solely on publicly available sources.

Where psychometric properties have been reported in more than one study/validation exercise, we will document this and appraise the psychometric performance of the measure based on the most recent UK standardisation exercise. Where evidence from a UK standardisation exercise is not available, we will base our appraisal on the most relevant instance (for example, administration with a nationwide sample in the UK context, administration in English, closest age group, largest sample size etc.). We will clearly document the rationale for the selected source of evidence. Where an EEF trial is the only available source, this will be used as the main evidence base. Where other evidence exists, EEF findings will be reported separately under a section on 'learning from use in EEF trials.' In such cases, we will note any issues encountered in practice, particularly the presence of floor or ceiling effects in large-scale use, and the contexts in which these arose (for example, specific age groups or settings).

We will carry out this appraisal for all new measures added to the database and for measures already included in the database. For existing measures, we will ensure the documented information is up to date and add any additional information (e.g., information based on a new norming exercise).

Table 2: Criteria for psychometric appraisal

Criterion (relevant pages in EFPA, 2013)	Factor assessed	Rating
Standardisation sample (p27, 35)	UK standardisation sample size	<p>N/A: Not standardised</p> <p>N/A: Standardised with a non-UK population</p> <p>✓: <200</p> <p>✓✓: 200-999</p> <p>✓✓✓: ≥1000</p>
	UK standardisation sample characteristics	<p>N/A: Not standardised</p> <p>N/A: Standardised with a non-UK population</p> <p>✓: No attempt was made to collect a random or stratified sample. Sample may be entirely opportunistic (e.g. convenience samples). Demographic or contextual variables (e.g. age, gender, ethnicity) are not reported or show clear imbalance relative to the target population.</p> <p>✓✓: Some attempt was made to recruit a balanced sample (e.g. stratification, quotas) but balance was not fully achieved. Sample may be broadly diverse but still underrepresents certain groups or demographics are not reported after random sampling model used.</p> <p>✓✓✓: Data were gathered by means of a random sampling model and good representativeness is established with respect to variables such as gender, age, ethnicity)</p>
Validity (p54-60)	Construct validity: how well a test measures the psychological construct it claims to assess.	<p>✓: No evidence that items and test were developed based on theory and construct definition OR statistical tests (e.g. Factor Analysis, Item-test correlations) demonstrate lack of construct validity (see EFPA guidance for further detail)</p> <p>✓✓: Evidence that items and test were developed based on theory and construct definition</p> <p>✓✓✓: Evidence that items and test were developed based on theory and construct definition AND statistical tests demonstrate construct validity confirm construct validity (see EFPA guidance for further detail)</p>
	Criterion validity: how well test scores correlate with relevant external criteria	<p>N/A: Not criterion referenced</p> <p>✓: Inadequate ($r < 0.20$) correlation with the criterion measure OR small (<100) sample size</p> <p>✓✓: Adequate ($0.20 \geq r < 0.50$) correlation with the criterion measure AND adequate sample size (≥ 100)</p> <p>✓✓✓: Excellent ($r \geq 0.50$) correlation with the criterion measure AND adequate sample size (≥ 100); OR Good ($0.35 \geq r < 0.50$) correlation with the criterion measure shown in one large sample (>200) or more than one adequate sized study</p>

Criterion (relevant pages in EFPA, 2013)	Factor assessed	Rating
Reliability (p43-51)	Internal consistency	<p>Not reported</p> <p>✓: Inadequate ($r < 0.70$) reliability coefficient OR small (<100) sample size</p> <p>✓✓: Adequate ($0.70 \leq r < 0.80$) reliability coefficient AND adequate sample size (≥ 100)</p> <p>✓✓✓: Excellent ($r \geq 0.90$) reliability coefficient AND adequate sample size (≥ 100); OR Good ($0.80 \leq r < 0.90$) reliability coefficient in one large sample (>200) or more than one adequate sized study</p>
	Test re-test reliability	<p>Not reported</p> <p>✓: Inadequate ($r < 0.60$) reliability coefficient OR small (<100) sample size</p> <p>✓✓: Adequate ($0.60 \leq r < 0.70$) reliability coefficient AND adequate sample size (≥ 100)</p> <p>✓✓✓: Excellent ($r \geq 0.80$) reliability coefficient AND adequate sample size (≥ 100); OR Good ($0.70 \leq r < 0.80$) reliability coefficient in one large sample (>200) or more than one adequate sized study</p>
	Inter-rater reliability	<p>Not reported</p> <p>✓: Inadequate ($r < 0.60$) reliability coefficient OR small (<100) sample size</p> <p>✓✓: Adequate ($0.60 \leq r < 0.70$) reliability coefficient AND adequate sample size (≥ 100)</p> <p>✓✓✓: Excellent ($r \geq 0.80$) reliability coefficient AND adequate sample size (≥ 100); OR Good ($0.70 \leq r < 0.80$) reliability coefficient in one large sample (>200) or more than one adequate sized study</p>

Documenting additional information

We will document additional information for each shortlisted measure (including both newly identified measures and those currently in the database), from test documentation and test implementation manuals. This will include:

- the skills, sub-domains and domains tested by the measure, as per the new EYMD taxonomy (**Error! Reference source not found.**)
- the original intended use of the assessment – e.g. developmental screening, outcome measurement in evaluation, progress tracking, population-level measurement⁶
- sector links, including relevant EYFS Early Learning Goals, relevant EYES approaches, and availability in national administrative datasets
- administration format (individual or group)
- respondent (child, parent, teaching professional, other)
- response mode (behavioural interaction; drawing; keyboard or mouse responses; manual (physical) operations; oral; paper and pencil; touch screen)⁷

⁶ These categories will be refined in liaison with EEF as we review measures.

⁷ This list is taken from the EFPA (2013) review model and may be refined as we document the response mode of shortlisted measures.

- item response type (Graded scale ratings (e.g., Likert); performance/task success; interactions; multiple choice/forced choice; open response; rankings; time (e.g., latency, reaction speed, completion time)⁸
- prior qualifications/training required to administer the measure
- the nature of scoring the measure, including whether expert judgement is required to score participants, and whether complex formulae are required
- availability of guidance for administering and scoring the measure
- time required to administer the measure
- cost and availability

We will also document any use of each measure in EEF-funded trials, as mentioned above. This will include a short description of the use and any key learnings. We will pay particular attention to issues with floor and ceiling effects for particular groups, and to any adaptations made and how the measure has performed in instances where it has been adapted for use in the UK context, or used with age ranges beyond which the measure was originally developed for/standardised with. This information will be extracted from published trial reports and may be supplemented with short interviews with evaluators if further information is required.

Presentation in a database

We will present the information outlined above in an Excel database, structured and coded/searchable by the final domains, sub-domains and skills, as well as the psychometric appraisal and other key features of the measures, such as target age group and mode and format of administration.

Deliverables

We will produce the following outputs from this strand of work:

- **Updated EYMD** as described above.
- **Final report** presenting findings from the review. This will include an outline of changes made to the EYMD and user interface; analysis of the availability of tools, strengths and weaknesses in measurement across the domains; an explanation of the adapted EFPA criteria for psychometric appraisal; and any other key methodological reflections.
- **Guidance on using the EYMD** to ensure EYMD can be used and updated by EEF and evaluators in the future. This will include an outline of the purpose of the EYMD and how to use it, and a selection of practical “top tips” on how to use the database effectively. It will include guidance on how the psychometric appraisal criteria should be interpreted, signposting to the final report with further information on how the EFPA criteria were adapted.
- **An extraction template and supporting guidance** for EEF to extract data for incorporating into the database from EEF-funded early years trials and other sources of information about relevant early years outcome measures.
- **A guidance note on adapting international measures for use in the UK**, including guidance on translation, cultural adaptation and validation processes.
- **A guidance note on choosing and using assessments for very young children (0-3 year olds).**

⁸ This list is taken from the EFPA (2013) review model and may be refined as we document the item response type of shortlisted measures.

Revising the EYMD design interface

Aims

The aims of this work strand support overall project objective 3 (see 5). The aims of the strand are twofold; first, to redesign the EYMD interface to improve its design usability and functionality; and second, to improve the rating system used to assess the quality and practicality of early years assessment tools.

Approach

Initial mock-up of database design

Building on initial scoping and insights from the EEF's internal review, we developed early design concepts for the revised EYMD interface. These included proposed changes to the layout, filtering functionality, and rating system. For the rating system, we tested a shift from the current star rating system, which rates the availability of information, to a more intuitive RAG (Red, Amber, Green) rating scheme across several key indicators, such as standardisation, validity, reliability, cost and availability, time for administration, scoring and required qualifications of the administrator. The revised system was designed to help users more quickly interpret technical information and identify strengths and weaknesses across multiple dimensions, making the database more useful and user-friendly

EEF consultation

We conducted a dedicated workshop with EEF staff, bringing together colleagues with experience in early years programmes and evaluations and with familiarity with the web design and technical structure of EEF's website and databases. The purpose of the workshop was to:

- Explore findings from EEF's internal review of the EYMD, identifying current limitations and opportunities for improvement.
- Discuss key design considerations and technical constraints, to ensure proposed changes are feasible within EEF's existing systems.
- Gather feedback on an initial mock-up of the revised EYMD interface, with a focus on usability and clarity.

Findings from the workshop informed revisions to the database mock-up and guided the next phase of user testing with evaluators. After consultations with EEF, we moved away from the RAG rating approach to the scoring described in Table 2 above. We decided that rating some criteria as "red" may suggest that particular measures should never be used, when in fact evaluators and researchers will always need to make informed choices about what is most important in judging the merits of available outcome measures, what compromises can be made, and why.

Evaluator survey

We invited a group of evaluators to participate in an online user experience survey. We worked closely with EEF to identify suitable participants, with a focus on those with experience working on early years evaluations or experience using the EEF's measures databases. We aimed to include evaluators familiar with the existing EYMD and those new to the platform.

The purpose of the survey was to gather broad, structured feedback on a mock-up of the revised EYMD interface, focusing on the usability, clarity, and perceived usefulness of the updated design and rating system. Participants were asked to feedback on specific features, such as filtering and domain navigation, and share their views on what worked well and where improvements were needed. The survey was conducted using the user research and usability testing platform Lyssna.

We received four responses to the survey. Feedback from the survey was analysed using a mixed-methods approach. Quantitative ratings were summarised and interpreted, while qualitative comments were reviewed individually,. The results were used to refine the database mock-up ahead of the next stage of consultation.

Consultation workshop with evaluators

Following the survey, we ran a workshop with three evaluators. We invited those invited to the survey to take part, as well as using a "snowball" approach asking them to pass the invite on to relevant colleagues.

This workshop built on the survey findings, using a revised version of the mock-up that incorporated the feedback received. The session explored evaluators' specific needs when identifying and selecting assessment tools, their challenges in navigating the current EYMD, and their preferences for how key psychometric and implementation information should be displayed and rated. The workshop also tested refined filtering options and design features to ensure they meet real-world use cases. We also sought feedback on the new rating system. Insights from this session will guide the final round of revisions and inform the final recommendations for the EYMD platform refresh.

Revision of database design mock-up

Findings from the workshops and evaluator survey will inform revisions to the mock-up. We may test the mock-up further with evaluators at this stage if deemed necessary. The refined design will reflect user preferences and practical needs, supporting both high-level exploration and in-depth appraisal of assessment tools. Final recommendations will be shared with the EEF for review and sign-off.

Deliverables

The outputs for this strand will be:

- EYMD website mock-up
- Supporting recommendations for the EEF's website team and designers

The EEF will use these outputs and the final database to update the EYMD platform and user interface.

Data protection

This project will not entail large scale data collection. However, it will involve collecting and processing personal data to facilitate consultations with EEF staff and sector experts. Data subjects will include EEF staff, early years sector experts and evaluators.

Personal data collected for this evaluation will only be used for the purpose of the project. We will store and handle data in line with the General Data Protection Regulation (GDPR) and Data Protection Act 2018. Only named individuals will have access to personal data and the team will comply with information security procedures that include preventative measures and processes for reporting, reviewing, and responding to breaches. We will securely delete personal data six months after the project ends (August 2026).

We will outline data protection procedures and safeguards and our legal bases for processing data in a privacy notice online and circulate it to all concerned parties. Our legal basis for data processing will be 'legitimate interests'.

We will not present specific findings from the consultations externally. When sharing findings with EEF, we will take steps to ensure that views are not attributable to individuals. We will communicate appropriate limits to confidentiality and anonymity to those consulted.

Risk management

Within our project management procedures, the Project Director and Project Manager have the responsibility of maintaining a risk register. The risk register will allow us to identify and communicate risks to EEF in a timely manner, as well as to develop mitigation strategies. The key risks identified for this project are outlined below in Table 3, including indication of their pre-mitigation/contingency likelihood and impact, and our mitigation and contingency strategies.

Table 3: Risk matrix

Risk	Likelihood	Impact	Mitigation/contingency
Project outputs do not meet EEF and users' needs	Low	High	<ul style="list-style-type: none"> Use inception phase to refine project aims, methodology, approaches and timelines. Decisions documented in project protocol. QA against protocol and agreed quality indicators. Agree with EEF at outset who will have ultimate sign-off for which elements of the project. Agree with EEF at outset the scope of the EYMD, based on EEF's organisational remit and strategic priorities. Engagement of EYMD users through workshop and survey.
Incomplete or inconsistent understanding of appraisal criteria leads to inconsistent application of screening criteria	Medium	High	<ul style="list-style-type: none"> Detailed briefing of the team on the criteria, scoring rubrics and how to escalate any uncertainties for discussion with the PI/senior advisors. Quality assurance procedures, including multiple researchers reviewing the same initial set of measures to check for consistency. At least three outcome measures coded by each member of the team will be double coded by two researchers. Discrepancies in application to be discussed and resolved before full screening begins. If large discrepancies remain, calibration will be repeated with further outcomes. Spot checks of screening decisions to be conducted by a senior member of the project team (PI or advisor). This will, initially be carried out on a sample of 10-20% of measures reviewed by each researcher, with the level of cross-checking subject to being adjusted to balance rigour with feasibility.
Duplication of existing efforts	Medium	Medium	<ul style="list-style-type: none"> Review of existing database at inception. In-depth consultation with EEF to understand work done by the organisation and UCL IOE during initial review of EY assessment measures to ensure this project builds on that (for example, focussing on identifying measures developed or updated from 2017 onwards for existing domains). Team with good knowledge and experience of existing EYMD.
Not feasible to appraise all long-listed measures within resources available	Low	Medium	<ul style="list-style-type: none"> Clearly defined research protocol and targeted search strategy to ensure the longlist remains focused, relevant, and manageable. Criteria for prioritisation of measures if needed agreed with EEF. Checkpoints with EEF to report progress and co-decide on thresholds or any trimming needed to remain within scope. Clear budget and timeline so that MeasurEd and EEF can work together to ensure we are on track. EYMD to be a live document that EEF can continue to update.
Delays due to resourcing gaps	Low	Medium	<ul style="list-style-type: none"> Committed team with full time allocation upfront. Use of established tools and processes to allocate, protect and adjust staff resource as needed. Wider pool of consultants with relevant skills and expertise to draw on if timelines or availability shift.
Poor engagement from evaluators in initial workshop and survey	Medium	High	<ul style="list-style-type: none"> Using EEF's known network to reduce cold contact. Clear value proposition, i.e., opportunity to shape a database evaluators can rely on. Including a variety of evaluators, e.g., independent evaluators and academics.

Risk	Likelihood	Impact	Mitigation/contingency
			<ul style="list-style-type: none"> Understanding power dynamics and creating safe spaces for discussions during workshops. Reduce time demand by limiting workshop time and keeping survey short. Strategies for engagement such as breakout sessions and visualisations.
Low acceptance of proposed rating system	Medium	Medium	<ul style="list-style-type: none"> Criteria for psychometric appraisal adapted from established (EFP)A criteria for psychometric appraisal. Clear explanation of these criteria in the protocol and final report. Clearly indicating the EFPA criteria used for rating on the EYMD interface (with approach to this informed by consultation with evaluators). Factors such as cost and time not rated due to variability across factors such as age and domain, and need to weigh up these practical factors against the information provided by the measure. Framing the ratings as a practical decision-support tool rather than a definitive judgement and including guidance on how users can interpret or supplement the ratings with their own professional judgement.
Technical or design limitations in mock-up	Low	Medium	<ul style="list-style-type: none"> Scope mock-up ambitions realistically. Use low-fidelity clickable prototypes. Test user experience rather than user interface.
Jingle-jangle fallacy leads to duplication or omission	High	Medium	<ul style="list-style-type: none"> Clear definition of domains and terminology. Report other known terminologies and definitions. Targeted search strategy using defined terminologies
Measures in database are not suitable for 0-3-year-olds or feasible to administer	Medium	High	<ul style="list-style-type: none"> Identify appropriate approaches to assessing 0-3-year-olds during review. Agree exclusion criteria with EEF around factors impacting feasibility such as duration. Include information in database about administration and factors identified. Draw on evaluator experiences in team to make decisions on suitability for inclusion. Transparent reporting of gaps in outcome measure availability – for example for this age range.
Few measures available for some domains	High	Medium	<ul style="list-style-type: none"> If necessary, expand the scope of search by using forward- and backward-citation tracking. If necessary, expand the scope of search to include measures published prior to 2017. EYMD to be a live resource over time, updated with learning from EEF trials and newly developed measures. Final report to summarise availability of measures across domains and ages, highlighting gaps for sector to address.
Information for appraisal not publicly available about measures	High	Low	<ul style="list-style-type: none"> Timeline to allow for contacting research teams who have used measures. Agree protocols at the outset with EEF about in/exclusion of measures with information missing. Flag gaps for EEF approval.

Team

Table 4 includes an overview of project team members, their role and institutional affiliation.

Table 4: Project team

Name	Role	Affiliation
Dr Lydia Marshall	Principal Investigator	Oxford MeasurEd (Director of Research)
Sara Bashir Malik	Project Manager and UX Lead	Oxford MeasurEd (Senior Consultant)
Anushay Mazar	Analyst, Early Years and SEL Specialist	Oxford MeasurEd (Consultant)
Astrid Pickenpack	Analyst	Oxford MeasurEd (Consultant)
Paulina Valenzuela	Analyst, Assessment Specialist	Oxford MeasurEd (Consultant)
Dr Rachel Outhred	Advisor	Oxford MeasurEd (Managing Director)
Robert Wishart	Advisor	Oxford MeasurEd (Associate)
Professor Abbie Raikes	Advisor	University of Nebraska (Professor)

Conflicts of interest

To our best knowledge, there are no actual or potential conflicts of interest regarding the project Update of the EEF Early Years Measures Database. Should any conflicts arise, we will report these promptly and initiate steps to manage or mitigate these in line with existing organisational policies, procedures as well as ethical standards.

Timeline

Table 5 presents a timeline of activities related to the review.

Table 5: Timeline

Dates	Activity	Staff responsible / leading
Reviewing the EYMD taxonomy and structure		
June 2025	Desk review of existing definitions	OM
July 2025	Rapid desk review of literature	OM
July – August 2025	Stakeholder consultation on taxonomy	OM
August 2025	Agree decisions on taxonomy with EEF	OM/EEF
August – September 2025	Further development of taxonomy	OM/EEF
July – September 2025	Writing and publishing protocol	OM/EEF
Updating the EYMD content and related information		
September – October 2025	Sourcing a longlist of outcome measures	OM
October 2025	Screening the longlist to create shortlist	OM
October 2025 – January 2026	Documenting additional information and presenting in a database	OM
November – December 2025	Interviews with evaluators to gather additional information (if needed)	OM
January 2026	Database spreadsheet finalised	OM
January – February 2026	Creating guidance on using EYMD, adapting international measures for use in the UK and choosing and using assessments with very young children	OM

Dates	Activity	Staff responsible / leading
January – March 2026	Writing and publishing report with findings and conclusion	OM/EEF
Revising the EYMD design interface		
June – July 2025	Initial mock-up of database design	OM
July 2025	Initial EYMD design and usability consultation workshop with EEF	OM/EEF
July 2025 –September 2025	Evaluator survey for user testing of initial mock-up	OM
September 2025	Consultation workshop with evaluators on the EYMD design and user experience	OM
September 2025	Agree on design decisions with EEF	OM/EEF
September – October 2025	Revision of database design mock-up and final recommendations	OM/EEF
Publication of the EYMD		
March 2026	Publication of the updated EYMD, final report and accompanying guidance	EEF

References

Britto, Pia R., et al., (2017) 'Nurturing Care: Promoting early childhood development', *The Lancet*, vol. 389, no. 10064, January, pp. 91–102, Available at: www.sciencedirect.com/science/article/pii/S0140673616313903

Department for Education (2021). 'Changes to the early years foundation stage (EYFS) framework'. 01 September 2021. Available at: <https://www.gov.uk/government/publications/changes-to-the-early-years-foundation-stage-eyfs-framework/changes-to-the-early-years-foundation-stage-eyfs-framework>

Department for Education (2025a). 'Changes to the EYFS framework'. Available at: <https://help-for-early-years-providers.education.gov.uk/support-for-practitioners/changes-to-the-eyfs-framework>

Department for Education (2025b). *Areas of Research Interest*. April 2025. Available at: https://assets.publishing.service.gov.uk/media/68136b1bb0ef2c985052541c/DfE_areas_of_research_interest.pdf

Dockrell, J., Hurry, J., Cowan, R., Flouri, E., & Dawson, A. (2017). Review of assessment measures in the early years: Language and literacy, numeracy and social emotional development and mental health. December 2017. Education Endowment Foundation: London. Available at: https://discovery.ucl.ac.uk/id/eprint/10039319/8/Dockrell_Report.pdf

EFPA (2013) EFPA Review Model for the Description and Evaluation of Psychological and Educational Tests. Test review form and notes for reviewers. Version 4.2.6.b Available at https://www.efpa.eu/sites/default/files/2023-06/110c_EFPA_BOA_TEST REVIEW_MODEL_version426.pdf⁹

Fernald, L. C., Prado, E., Kariger, P., & Raikes, A. (2017). A Toolkit for Measuring Early Childhood Development in Low and Middle-Income Countries. *World Bank Publications*.

Hodgen, J., Barclay, N., Foster, C., Gilmore, C., Marks, R., & Simms, V. (2020). Early Years and Key Stage 1 Mathematics Teaching: Evidence Review. Education Endowment Foundation. https://educationendowmentfoundation.org.uk/public/files/Early_Years_and_Key_Stage_1_Mathematics_Evidence_Review.pdf

Law, J., Charlton, J., Dockrell, J., Gascoigne, M., McKean, C., & Theakston, A. (2017). Early Language Development: Needs, provision, and intervention for preschool children from socio-economically disadvantaged backgrounds. *London: Education Endowment Foundation*, 204.

Tierney, A. L., & Nelson III, C. A. (2009). Brain development and the role of experience in the early years. *Zero to three*, 30(2), 9.

⁹ The EFPA Test Review Criteria were largely modelled on the form and content of the British Psychological Society's (BPS) test review criteria and criteria developed by the Dutch Committee on Tests and Testing (COTAN) of the Dutch Association of Psychologists (NIP). EFPA is grateful to the BPS and NIP for permission to build on their criteria in developing the European model. All intellectual rights in the original BPS and NIP criteria are acknowledged and remain with those bodies.

Annex A: Draft taxonomy

Domain	Domain Definition	Sub-Domain	Sub-Domain Definition	Skill	Skill Definition
Communication and Language	Children's proficiency in understanding and conveying information, ideas and feelings, including through words and sentences.	Communication	Children's proficiency in conveying and understanding information, ideas or feelings	Joint attention	Proficiency in sharing focus with one or more individuals on the same object or event, coordinated through eye gaze, pointing, or other verbal/non-verbal cues.
				Gesticulation	Proficiency in using gestures, such as hand movements or facial expressions, to convey meaning alongside or instead of speech.
				Turn taking	Proficiency in engaging in reciprocal communication with another person.
				Speech	Proficiency in producing spoken language sounds clearly and fluently to convey meaning.
				Pragmatic skills	Proficiency in adapting verbal and non-verbal communication to different social contexts.
				Code switching	Proficiency in alternating between two or more languages or language varieties depending on context and audience.
		Language	Children's proficiency in understanding words and sentence structures and using them to express meaning.	Expressive vocabulary	The range of words a child can produce and use in speech, writing or sign language.
				Receptive vocabulary	The range of words a child can understand when heard or read.
				Syntax	Proficiency in arranging words to make meaning
				Morphology	Proficiency in using word parts, such as prefixes and suffixes, to create meaning.
				Grammar	Proficiency in using the system of rules governing the structure of words, phrases and sentences, to create meaning.
				Narrative skills	Proficiency in structuring and telling a story or recounting events coherently.
				Inferencing	Proficiency in understanding information that is implied but not directly stated, using context clues and prior knowledge.
				Language reasoning	Proficiency in using language to explain their thinking and reasoning.

Domain	Domain Definition	Sub-Domain	Sub-Domain Definition	Skill	Skill Definition
Literacy	Children's proficiency in engaging with, understanding and conveying meaning through written text	Reading	Children's proficiency in decoding and making meaning from written text.	Print awareness	Understanding that print carries meaning and knowing how to handle books and written materials.
				Letter naming	Proficiency in recognising and naming the letters of the alphabet.
				Phonological awareness	Proficiency in recognising and manipulating the sound structures of language (e.g. syllables, phonemes, rhyme and alliteration).
				Letter-sound correspondence	Proficiency in matching written letters/graphemes and the sounds they represent.
				Word reading	Proficiency in reading familiar and unfamiliar words.
				Rapid automised naming	Proficiency in quickly naming familiar items, such as letters, words, numbers or colours, when presented visually.
				Code switching	Proficiency in alternating between two or more languages or language varieties depending on context and audience.
				Reading comprehension	Proficiency in understanding and interpreting the meaning of written text.
		Writing	Children's proficiency in generating and recording ideas using symbols to communicate meaning.	Reading fluency	Proficiency in reading text accurately, quickly, and with appropriate expression.
				Graphomotor skills	Proficiency in coordinating fine motor movements for writing and drawing.
				Mark making	Proficiency in creating marks on a surface as a precursor to writing.
				Spelling	Proficiency in writing words with correct letter sequences to represent spoken language.
				Handwriting	Proficiency in forming letters and words legibly and consistently.
				Written composition	Proficiency in organising ideas to create meaningful written text.

Domain	Domain Definition	Sub-Domain	Sub-Domain Definition	Skill	Skill Definition
Mathematics	Children's proficiency in engaging with and understanding ideas about number, relationships, quantity and space.	Number	Children's proficiency in understanding and expressing ideas about quantity, counting and numerical patterns.	Pattern awareness/matching	Proficiency in recognising and reproducing patterns using shapes, colours, numbers, objects or events.
				Recognising and naming digits	Proficiency in identifying and naming written numbers.
				Linking number words, digits, and quantities	Proficiency in identifying the relationship between spoken number words, written digits, and their quantities.
				Counting	Proficiency in skills including correctly saying number words in a sequence, assigning each number word to a single object (one-to-one correspondence), and recognising that the final number word used represents the total number of objects in the set (cardinality).
				Subitising	Proficiency in visually seeing a number of objects instantly without needing to count them out one at a time.
				Number composition	Understanding that a number can be made up of two or more smaller numbers.
				Place value	Understanding the value of a digit based on its position within a number.
				Comparing numerical magnitude	Proficiency in identifying which numbers are worth more or less than each other.
		Operations	Children's proficiency in grouping, sharing and combining numbers.	Numerical patterns	Proficiency in recognising and reproducing patterns using numbers, and what this means about their relationships.
				Mathematical talk	Proficiency in having spoken or written conversation with teachers, children and self in order to problem solve or aid understanding of mathematical concepts.
		Operations	Children's proficiency in grouping, sharing and combining numbers.	Additive reasoning	Proficiency in solving problems involving combining, separating, or comparing quantities using addition and subtraction.
				Calculation	Proficiency in carrying out mathematical operations such as addition, subtraction, multiplication, or division.
				Inverse operations	Understanding that some mathematical operations can reverse the effect of others.
				Symbol knowledge	Proficiency in recognising and understanding mathematical symbols and their meanings.
				Strategy selection	Proficiency in choosing appropriate methods to solve mathematical problems.
				Comparing numerical magnitude	Proficiency in identifying which numbers are worth more or less than each other.
				Numerical patterns	Proficiency in recognising and reproducing patterns using numbers, and what this means about their relationships.

Domain	Domain Definition	Sub-Domain	Sub-Domain Definition	Skill	Skill Definition
				Mathematical talk	Proficiency in having spoken or written conversation with teachers, children and self in order to problem solve or aid understanding of mathematical concepts.
				Non-standard units	Proficiency in measuring objects or quantities with everyday items rather than formal units, such as using blocks, footsteps, or hand spans
				Estimation	Proficiency in making a reasonable guess about quantity, size, or outcome before measuring or counting
				Comparing physical magnitude	Proficiency in comparing and ordering objects and shapes by size, weight and capacity
				Comparing quantities	Proficiency in identifying which quantities are worth more or less than each other.
				Pattern awareness/matching	Proficiency in recognising and reproducing patterns using shapes, colours, numbers, objects or events.
				Mathematical talk	Proficiency in having spoken or written conversation with teachers, children and self in order to problem solve or aid understanding of mathematical concepts.
				Shape recognition	Proficiency in identifying and naming common shapes.
				Shape decomposition	Understanding that complex shapes can be broken down into smaller, simpler shapes, and ability to do so.
				2D-3D relationships	Understanding how two-dimensional representations relate to three-dimensional objects.
				Spatial language	Proficiency in using words to describe the position, direction, and movement of objects in space.
				Perspective taking	Proficiency in understanding and representing how objects or scenes appear from different viewpoints.
				Comparing physical magnitude	Proficiency in comparing and ordering objects and shapes by size, weight and capacity
				Pattern awareness/matching	Proficiency in recognising and reproducing patterns using shapes, colours, numbers, objects or events.
				Mathematical talk	Proficiency in having spoken or written conversation with teachers, children and self in order to problem solve or aid understanding of mathematical concepts.

Domain	Domain Definition	Sub-Domain	Sub-Domain Definition	Skill	Skill Definition
Personal, Social, and Emotional Development	Children's proficiency in forming relationships, understanding themselves and making choices about their emotions, behaviour and learning.	Personal Development	Children's understanding of themselves and capacity to take actions to look after themselves.	Self-esteem	Confidence in their own worth or abilities.
				Self-concept/ sense of self	Beliefs and knowledge about their own attributes and qualities
				Self-care and healthy behaviours	Capacity to engage in activities and habits that maintain physical and mental health, such as rest, hygiene, and healthy eating.
				Safety behaviours	Capacity to engage in activities and habits to avoid harm and ensure personal safety.
		Social Skills	Children's proficiency in adapting to and interacting effectively with others in group and one-to-one contexts.	Theory of mind/ mentalisation	Understanding that others have thoughts, feelings, and perspectives different to their own.
				Empathy	Proficiency in understanding and sharing others' feelings.
				Relationship skills	Proficiency in interacting effectively with others, building positive relationships, and managing social challenges.
				Prosocial behaviour	Capacity to engage in actions intended to benefit others, such as helping, supporting or sharing.
				Conflict resolution skills	Proficiency in resolving disagreements in respectful and constructive ways.
		Emotion Skills	Children's proficiency in recognising, understanding and expressing emotions in themselves and others.	Emotion recognition	Proficiency in identifying their own emotions through feelings and through verbal and non-verbal communication.
				Emotion expression	Proficiency in communicating feelings verbally and nonverbally.
				Emotion vocabulary	The range of words a child can use to describe and express feelings and emotions.
				Emotion understanding	Proficiency in understanding a range of distinct emotions that they can experience.
				Emotional regulation	Proficiency in planning, monitoring, and adjusting their emotional responses.
		Self-Regulation	Children's proficiency in adapting their thoughts, behaviours, and emotions to achieve their goals.	Emotional regulation	Proficiency in planning, monitor, and adjust their emotional responses.
				Behavioural regulation	Proficiency in planning, monitoring, and adjusting their actions and behaviours.
				Cognitive regulation	Proficiency in planning, monitoring, and adjusting their thinking processes to achieve goals.
				Persistence	Proficiency in continuing working toward a goal despite challenges.
				Inhibition	Proficiency in controlling impulses and refrain from action.
		Executive Functioning	Children's proficiency in using cognitive skills to carry out goal-directed behaviour.	Meta-cognition	Proficiency in thinking about and reflecting on their own thinking and learning processes.
				Cognitive flexibility	Proficiency in adjusting to situational demands, shifting perspectives, and balancing competing needs.

Domain	Domain Definition	Sub-Domain	Sub-Domain Definition	Skill	Skill Definition
				Working memory	Proficiency in holding and manipulating information in mind for short periods.
Physical Development	Children's proficiency in the physical skills needed for everyday tasks and active play.	Gross Motor Skills	Children's proficiency in coordinating their large muscles to balance and move around.	Attention control	Proficiency in focusing attention on relevant information while ignoring distractions.
				Persistence	Proficiency in continuing working toward a goal despite challenges.
				Inhibition	Proficiency in controlling impulses and refrain from action.
				Stability/ balance	Proficiency in maintaining control of their body position whether moving or still.
				Coordination	Proficiency in using different parts of their body together smoothly and efficiently.
				Body awareness	Understanding where their body is in space and how it moves.
		Fine Motor Skills	Children's proficiency in coordinating the small muscles in their hands and fingers to perform precise tasks.	Strength	Proficiency in exerting force to move objects or their body.
				Locomotion	Proficiency in moving from one place to another using actions such as crawling, walking, running, or jumping.
				Bilateral coordination	Proficiency in coordinating both sides of their body to perform tasks.
				Grasping and releasing	Proficiency in grasping and letting go of objects intentionally.
				In-hand manipulation	Proficiency in moving and positioning objects within one hand.
				Hand-eye coordination	Proficiency in using visual information to guide their hand movements.
				Palmar grasp	Proficiency in holding objects securely using the whole hand, with the fingers wrapped around the object
				Pincer grasp	Proficiency in holding small objects between their thumb and forefinger.
				Graphomotor skills	Proficiency in coordinating fine motor movements for writing and drawing.
				Bilateral coordination	Proficiency in coordinating both sides of their body to perform tasks.

Domain	Domain Definition	Sub-Domain	Sub-Domain Definition	Skill	Skill Definition
Expressive Arts & Design	Children's proficiency in exploring and expressing their ideas, feelings and creativity.	Performing Arts	Children's proficiency in exploring and expressing ideas, feelings and creativity through movement, music and drama.	Early musicality	Proficiency in perceiving, producing, and responding to musical elements such as rhythm and pitch.
				Role play	Proficiency in taking on and acting out roles in imaginative scenarios.
				Narrative	Proficiency in structuring and telling a story or recounting events coherently.
				Dance	Proficiency in using movement intentionally to express ideas, emotions, or stories.
				Creativity	Proficiency in generating and expressing original ideas.
		Visual Arts	Children's proficiency in exploring and expressing ideas, feelings and creativity through drawing, painting and other visual media.	Colour knowledge	Proficiency in identifying different colours and their properties.
				Manipulating materials	Proficiency in handling and shaping materials to create or modify objects.
				Creativity	Proficiency in generating and expressing original ideas.
		Design & technology	Children's proficiency in creating and adapting materials, structures and tools to express ideas and solve problems.	Understanding function	Understanding how tools, machines, or objects help people in daily life.
				Selecting materials	Proficiency in choosing appropriate materials or tools for an identified problem or purpose.
				Manipulation materials	Proficiency in handling and shaping materials to create or modify objects.
				Creativity	Proficiency in generating and expressing original ideas.
Understanding the World	Children's proficiency in making sense of the social and material world around them.	Past & Present	Children's proficiency in engaging with and understanding ideas about the order of events and the passage of time.	Understanding of sequence	Understanding of the order of events or steps in a process.
				Chronological understanding	Understanding the passage of time and how events are ordered in relation to each other.
				Chronological vocabulary	The range of words a child can use to describe time and sequence.
				Comparison skills	Proficiency in identifying similarities and differences between objects, concepts, or events.
				Observation skills	Proficiency in closely watching and noticing details in the environment.
		The Natural World	Children's proficiency in engaging with and understanding ideas about living things and the material environment.	Comparison skills	Proficiency in identifying similarities and differences between objects, concepts, or events.
				Observation skills	Proficiency in closely watching and noticing details in the environment.
				Natural process knowledge	Knowledge of phenomena related to nature and the environment.
		People, Culture & Communities	Children's proficiency in engaging with, understanding and appreciating the social world.	Comparison skills	Proficiency in identifying similarities and differences between objects, concepts, or events.
				Observation skills	Proficiency in closely watching and noticing details in the environment.

Domain	Domain Definition	Sub-Domain	Sub-Domain Definition	Skill	Skill Definition
				Religious and cultural knowledge	Knowledge of different religious and cultural beliefs, practices and traditions.
				Sensitivity to difference	Awareness and respect for individual and cultural differences.
Digital Technology ¹⁰	Children's understanding of proficiency in using digital tools and devices.	Technology Awareness	Children's understanding of the purposes and functions of digital technology.		
		Digital Literacy	Children's proficiency in using digital tools to find, create and communicate information and solve problems.		
Early Learning Environment	The physical, social, and emotional conditions that shape children's early development and learning.	Home Learning Environment	The physical, social, and emotional conditions that shape children's early development and learning at home.	Physical environment	The physical surroundings where learning and development take place.
				Social and emotional climate	The relationships, emotional support, and interactions in a learning environment.
				Adult-child interactions	Adults' interactions with children.
				Peer interactions	Children's interactions with each other.
				Health and hygiene routines	Children's understanding and practise of everyday habits that keep the body clean, healthy, and safe.
				Learning resources	The materials and tools available to support children's learning and development..
				Cognitive stimulation	The experiences and activities introduced to support children's learning and problem-solving.
	The physical, social, and emotional conditions that shape children's early development and learning in a non-home environment (e.g. formal early childhood education settings, outdoor learning provision).	Early Years Setting Learning Environment	The physical, social, and emotional conditions that shape children's early development and learning in a non-home environment (e.g. formal early childhood education settings, outdoor learning provision).	Physical environment	The physical surroundings where learning and development take place.
				Social and emotional climate	The relationships, emotional support, and interactions in a learning environment.
				Adult-child interactions	Adults' interactions with children.
				Peer interactions	Children's interactions with each other.
				Health and hygiene practices	Health, safety and hygiene routines and protocols.
				Learning resources	The materials and tools available to support children's learning and development.
				Cognitive stimulation	The experiences and activities introduced to support children's learning and problem-solving.

¹⁰ Included as a set of skills underpinning learning and development in other domains.

Annex B: Search terms

Population Keywords

(child* OR toddler* OR infant* OR preschool* OR pre-school* OR "early childhood" OR "young children" OR "ECE" OR "ECEC" OR "early years" OR "nursery" OR "kindergarten" OR "reception" OR "baby room" OR "pre primary" or pre-primary)

Measurement/Assessment Keywords

(measure* OR tool* OR instrument* OR assessment* OR test* OR scale* OR checklist* OR observation*)

Domain-Specific Keywords

Language and Communication

(language OR communication OR vocabulary OR grammar OR syntax OR morphology OR speech OR "turn taking" OR "pragmatic skills" OR "pragmatic language" OR "joint attention" OR "shared attention" OR gestur* OR "gesticulation" OR "narrative skills" OR "storytelling" OR "narrative competence" OR inferenc* OR "language reasoning" OR "verbal reasoning" OR "code switching" OR "language switching" OR "bilingual language use" OR "EAL")

Literacy

("literacy" OR "pre-literacy" OR "reading" OR "writing" OR "print awareness" OR "print concepts" OR "concepts of print" OR "print knowledge" OR "emergent literacy" OR "book handling skills" OR "understanding of print" OR "text directionality" OR "recognition of print" OR "print recognition" OR "letter naming" OR "letter name knowledge" OR "alphabet knowledge" OR "letter identification" OR "letter recognition" OR "naming letters" OR "phonological awareness" OR phoneme OR phonics OR "phonological processing" OR "sound awareness" OR "sound discrimination" OR "syllable segmentation" OR "rhyme" OR "letter-sound" OR "grapheme-phoneme correspondence" OR "sound-symbol knowledge" OR alphabe* OR decoding OR "early word recognition" OR "rapid automatized naming" OR "rapid automised naming" OR "RAN" OR "rapid naming" OR "naming speed" OR "rapid serial naming" OR "naming fluency" OR "code switching" OR "language switching" OR "bilingual language use" OR "language alternation" OR "code mixing" OR translanguaging OR "comprehension" OR "prosody")

Mathematics

(math OR maths OR mathematics OR numeracy OR pre-numeracy OR "pattern awareness" OR "pattern matching" OR "digit recognition" OR "naming digits" OR "number word*" OR counting OR "one-to-one correspondence" OR cardinality OR subitising OR "number composition" OR "place value" OR "comparing *magnitude" OR "numerical patterns" OR "additive reasoning" OR calculation OR "inverse operations" OR "symbol knowledge" OR "strategy selection" OR "measur*" OR "non-standard units" OR "estimate*" OR "sequenc*" OR "shape recognition" OR "shape decomposition" OR "2D-3D relationships" OR "spatial language" OR "perspective taking")

PSED

("personal development" OR socioemotion* OR socio-emotion* OR self-esteem OR self-concept OR "sense of self" OR self-care OR "healthy behaviours" OR "safety behaviours" OR "social skills" OR "theory of mind" OR mentalis* OR empathy OR "relationship skills" OR "prosocial behaviour" OR "conflict resolution skills" OR "emotion skills" OR "emotion recognition" OR "emotion expression" OR "emotion vocabulary" OR "emotion naming" OR "emotion understanding" OR "emotion regulation" OR "emotional regulation" OR "self-regulation" OR "behavioural regulation" OR "behaviour regulation" OR "behavioral regulation" OR "behavior regulation" OR "cognitive regulation" OR persistence OR inhibit* OR "executive function*" OR metacognition OR "cognitive flexibility" OR "working memory" OR "attention control" OR "sustained attention" OR "task initiation")

Physical development¹¹

("gross motor skills" OR stability OR balance OR coordination OR "body awareness" OR strength OR locomotion OR "fine motor skills" OR grasping OR releasing OR "in-hand manipulation" OR "hand-eye coordination" OR "hand eye coordination" OR "palmar grasp" OR "pincer grasp" OR "graphomotor skills" OR "tool manipulation")

Expressive arts and design

(art* OR design* OR "design and technology" OR music* OR "role play" OR drama OR dance OR creativ* OR imaginat* OR sing OR "dramatic play" OR "colour knowledge" OR "manipulation of materials" OR sculpting OR paint* OR draw*)

Understanding the World

("understanding the world" OR "past and present" OR "order of events" OR "passage of time" OR "understanding of sequence" OR "chronological understanding" OR "chronological vocabulary" OR chronology OR "comparison skills" OR "observation skills" OR "the natural world" OR nature OR biology OR science OR "living things" OR "material environment" OR natur* OR geography OR "people, culture and communities" OR "religious and cultural knowledge" OR "religious education" OR "cultural education" OR "sensitivity to difference" OR "respect for diversity" OR "appreciation of differences" OR "cultural sensitivity" OR "cultural awareness" OR inclusivity OR open-mindedness OR "acceptance of others" OR "intercultural understanding" OR "social awareness" OR "valuing diversity" OR "recognition of individual differences" OR "awareness of social and cultural norms" OR "understanding of diverse perspectives")

Digital Technology

("technology awareness" OR "digital literacy" OR "technology skills" OR "ICT" OR "computing skills" OR "digital skills" OR "media literacy" OR "online literacy" OR "technological fluency" OR "digital competence" OR "digital awareness" OR "tech literacy" OR "information literacy" OR "computer literacy" OR "digital understanding" OR "digital navigation skills")

Early Learning Environments

("learning environment*" OR "learning resource*" OR "learning process quality" OR "adult-child interaction*" OR "teacher-child interaction*" OR "caregiver-child interaction*" OR "peer interaction*" OR "health routine*" OR "hygiene routine*" OR "care routine*" OR "cognitive stimulation" OR "teaching process quality" OR "instructional quality" OR "teacher quality" OR "teaching effectiveness" OR "classroom quality" OR "educational environment" OR "teacher-student interaction" OR "learning support" OR "early childhood education quality" OR "curricul*" OR "pedagog*")

¹¹ We have not included physical development as a search term as we believe the scope will be too broad. We will revisit this if the more specific search terms yield insufficient results. We will outline this or any other changes made to the search terms following the publication of this protocol, in the final report presenting findings from the review

You may re-use this document/publication (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence v3.0.

To view this licence, visit <https://nationalarchives.gov.uk/doc/open-government-licence/version/3> or email: psi@nationalarchives.gsi.gov.uk

Where we have identified any third-party copyright information you will need to obtain permission from the copyright holders concerned. The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

This document is available for download at <https://educationendowmentfoundation.org.uk>



**Education
Endowment
Foundation**

The Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
London
SW1P 4QP

<https://educationendowmentfoundation.org.uk>

X @EducEndowFoundn

f Facebook.com/EducEndowFoundn