

Abstract

Recently, the Children and Families Act 2014 was introduced in England to regulate provision for children with disabilities. According to this policy, statements of special educational needs were replaced with education, health and care plans, which should include high-quality, holistic and participation-focused outcomes to regulate provision; this change aligns with international recommendations regarding provision for children with disabilities. This study aimed to evaluate the outcomes defined for children with education, health and care plans in England.

236 Education Health and Care plans were included in the analysis, providing 2813 outcomes to be examined, which came from 11 local authorities and 42 schools and belong to 69 girls and 167 boys from 4 to 21 years of age.

The outcomes were independently rated by two experienced researchers using a Goal Functionality Scale. Inter-rater agreement was calculated for 10% of the outcomes.

Most outcomes were considered not to be functional or high-quality; differences in quality were found between local authorities, types of school, type of outcome, and the children's main need.

There are important quality concerns regarding the outcomes that have been designed for children with disabilities in England, which should be addressed through standardised training and guidelines on procedures.

What this paper adds

This is the first paper to date providing a systematic evaluation of outcomes designed for children with special educational needs and disabilities that have education, health and care (EHC) plans in England. Such evaluation is timely and of global importance for two reasons: first, the UK government has recently changed its policy for the provision of supports and services for children with disabilities (UK government, 2014), by replacing the statements of special educational needs with EHC plans, which should be holistic, participation-focused and contain SMART outcomes. However this has not been implemented without controversy, with professionals from various backgrounds expressing that the ideology of the new policy, although desirable, it is not being put into practice in the most effective way (Palikara, Castro, Gaona & Eirinaki, 2018); this paper provides evidence on the specific aspects of

current outcome design that need to see some improvement, with recommendations for policy and practice. Second, this paper highlights issues concerning the training of those involved in the Education Health and Care process, which are beyond the change in policy itself, but highlight the lack of alignment with internationally-recognised models of goal-setting and intervention for children with disabilities – for example, evidence is provided on the fact that most outcomes designed for children are not based on the current international definition of participation as ‘involvement in life situations’ (WHO, 2007). These two main sources of evidence provided by this paper question the compliance of the current English policy for disability provision with international conventions and call for specific quality assurance measures.

1. Introduction

1.1. Participation-focused provision

In countries with special education provision, children with Special Educational Needs and Disabilities (SEND) often have a legal document describing their needs, the provision to address those needs and the outcomes or goals resulting from that provision. It is through these documents that professionals, parents and others can track achievements and areas for improvement regarding attainment, but also regarding functioning in day-to-day life, over time. If disability is now understood as a restriction in *functioning* (WHO, 2001; 2007), then this should be regarded as the main objective of inclusive provision. The World Health Organisation (WHO, 2007) has defined *functioning* as the umbrella term for all aspects of participation, including activities performed, body functions and structures, and environmental factors (WHO, 2001; 2007), and *participation* as ‘involvement in life situations’ (p.248). These definitions have highlighted the role of natural life situations (understood as daily routines) as the context for SEND provision, instead of context-removed one-to-one interventions, to achieve full inclusion (Adolfsson, 2013). Therefore, focusing on daily life situations is adopting a participation-based approach. This contemporary view of SEND provision as *participation*-focused has been widely acknowledged for its alignment with United Nations’ conventions and for its fairer approach to disability, more focused on levels of individual functioning than on diagnostic labels (Simeonsson, 2009; Bornman, 2017); indeed it has been demonstrated that children’s needs are more dependent on their level of functioning than on their diagnosis (Castro & Pinto, 2015), and that participation levels have a predominant role in explaining development and learning, rather than the diagnostic labels (Pinto, Grande, Coelho, Castro, Granlund, & Björck-Åkesson, 2018). Similarly, it is widely accepted that two children with the same diagnosis often have very different functioning profiles (Simeonsson, Scarborough & Hebbeler, 2006). Therefore, devising *functional* outcomes for children (i.e., focusing on improving *participation* outcomes in daily routines) implies a definite move away from medical approaches to disability, positioning SEND provision beyond the social model of disabilities, in what can be designated as *functional* model (Simeonsson, 2006; Castro & Palikara, 2016). In many countries, this new functioning and participation-focused ideology has become the framework for national policy documents, where provision and often eligibility to services, are dependent on the individual participation profile, rather than on a specific diagnosis (e.g. Germany, Japan, Switzerland) (Sakai, Tanaka, & Tokunaga, 2017; Hollenweger, 2017;

Pretis, 2017). In England, recent policy changes have also shifted the SEND provision towards a more participation-focused approach by, for example, regulating what the outcomes designed for children should represent (Department for Education, 2015).

1.2 The English context and the link to an international participation-focused ideology for inclusion

Following the international trend towards participation-focused provision, recent policy changes in England entailed a shift from Statements of Special Educational Needs (which were mostly based on diagnosis) to Education Health and Care (EHC) plans (UK government, 2014). The EHC plans are the documents supporting the description of needs and respective provision for children with SEND in England. According to the Children and Families Act 2014, the EHC plans should be holistic, involving all areas of life, while still being individualised documents; the main aim of the EHC plan is to provide a clear picture of the child's functioning (including the child's own perspective) and behaviour and promote *participation*, which is explicitly mentioned as the main goal for provision in the Children and Families Act 2014 (e.g. p.21, 26, 57); this includes a tight collaboration and articulation between education, health and social care settings. This shift is well aligned with contemporary international literature in the field of inclusion and special educational needs, where participation has been regarded as a key descriptor for inclusion (e.g. McKay, Mahon, Donellan, Haracz, Sheldon & Ryan, 2017). Although a good body of literature has been growing on definitions of the *participation* concept, it is consensual that it should be understood as both *attendance* and *involvement* in life situations (Imms, Adair, Keen, Ullenhag, Rosenbaum & Granlund (2016), thus providing a very accurate indicator of inclusion.

Under Regulation 12 of the Special Educational Needs and Disability Regulations included in the Code of Practice accompanying the English Children and Families Act 2014, the EHC plans should have 11 sections: Section A should provide an overview of the views, interests and aspirations of the child/young person, and/or their parents; section B should provide a description of the child's education needs; section C should contain a description of the health needs of the child; section D should describe the social care needs of the child; section E refers to the outcomes sought for the child; section F should include the education provision that has been agreed in order to achieve the outcomes designed in the previous section; section G should refer to the health provision agreed; section H should refer to the

social care provision agreed; section I should specify the name of the school or other institution to be attended by the child; section J should detail the funding that might be made available for the child's provision; and finally section K should include copies of any advice, reports and assessment previously obtained as part of the overall education, health and care assessment procedure. In the current paper, we focus specifically on the section concerning the outcomes sought for the child/young person (section E), given the reported lack of clarity in relation to how these should be designed (Palikara, Castro, Gaona & Eirinaki, 2018). According to the SEND Code of Practice 2014, '*An outcome can be defined as the benefit or difference made to an individual as a result of an intervention. It should be personal and not expressed from a service perspective; it should be something that those involved have control and influence over, and while it does not always have to be formal or accredited, it should be specific, measurable, achievable, realistic and time bound (SMART)*'. (p.160).

The Code provides further guidance on outcomes clearly stating that the plans can include broader outcomes, such as long-term aspirations of the child and family, but these must not be included in the outcome section; outcomes should express '*the benefit or difference made to an individual as a result of an intervention*' (p. 163). However, the same Code of Practice also states that an outcome for a secondary school aged child can be, for example, '*to make sufficient progress or achieve a qualification to enable him or her to attend a specific course at college*' (p.163), which might not be considered SMART, and potentially not *functional*, as it does not focus on *participation* (defined as both *attendance* and *involvement* in daily life situations (Imms, Adair, Keen, Ullenhag, Rosenbaum & Granlund, 2016), but rather on attainment. While some flexibility regarding the design of the outcomes has been granted (the Code of Practice wording is that they *should* be SMART, but they do not necessarily *must* adopt that format), this caused controversy and meant that although professionals agree, in principle, with the ideology of the new policy for SEND provision, they don't think it has been adequately implemented for lack of more specific guidelines (Palikara, Castro, Gaona & Eirinaki, 2018), and for failing to provide a coherent picture involving education, health and social care (Boesley & Crane, 2018).

1.3 Defining quality outcomes

The literature on SMART outcomes has provided a number of definitions for each letter of the acronym over the years; It is consensual though, that SMART outcomes should be specific (e.g. Siegert & Taylor, 2004; Jung, 2007), measurable (Conzemius & O'Neill,

2011), action-oriented or attainable (e.g. Piskurich, 2015), relevant or realistic (e.g. Siegert & Taylor, 2004; Jung, 2007) and time-framed (e.g. Conzemius & O'Neill, 2011), which is why they are often considered of higher quality than outcomes that are not SMART. The relevance of the SMART criteria can be understood when considering international literature on what constitutes a quality outcome and how that is closely aligned with the definition of SMART outcomes. Table 1 provides an overview of the match between criteria used in the literature to define a good quality outcome and the definitions of the SMART acronym. For example, McWilliam (2009) proposes a set of criteria for *functional* goals in special education and early intervention (focused on participation in daily life routines): a) they should specify a daily routine in which the child will participate; b) they should specify precisely what the child is expected to do; c) the specified behaviour or skill should be relevant for the child's overall participation; d) there should be an acquisition criterion (when do we know that the child has acquired the behaviour?); e) the acquisition criterion should be relevant for the child's participation, i.e., embedded in a routine; f) there should be a generalization criterion (i.e. to other contexts and situations); g) there should be a timeframe for the acquisition of the new skill or behaviour. Based on these criteria, the Goal Functionality Scale (GFS) was developed. We argue that criteria a) and b) contribute for the outcome to be *specific*; criteria c) and e) relate to the *relevance* of the outcome; criteria d) and f) contribute to the *measurability* of the outcome; criterion e) relates to how *attainable* the outcome is; and criterion g) overlaps with the *timeframe* definition of the SMART acronym. Using these criteria and the GFS, Rakap (2015) looked at the quality of the outcomes written in the individualized education plans (IEPs) of 100 preschool children with disabilities in Turkey, concluding that the majority of the outcomes were of low quality, i.e., not functional or SMART. Boavida, Aguiar and McWilliam (2014) also looked at the goals in the IEPs of 83 preschoolers with disabilities in Portugal and rated them using the GFS. Results showed that IEP goals were too broad, not functional and measurable, and did not appropriately address skills within the context of natural routines and settings, i.e. *participation* (see table 1).

Other frameworks have emphasized the importance of functional outcomes for effective special education provision; for example, the embedded instruction model suggests that children with SEND need to have learning opportunities in their natural environments in order to acquire specific skills effectively and across contexts (Snyder, Rakap, Hemmeter, McLaughlin, Sandall, & McLean, 2015); within this framework, outcomes defined for children with SEND should specify who the learner is, the behaviour we are looking for, the

conditions in which the behaviour should occur, the activities performed to achieve that behaviour or skill and the criterion for acquisition (VanDerHeyden, Snyder, Smith, Sevin & Longwell (2005), which matches some of the SMART acronym definitions (see table 1); this makes outcomes developmentally-appropriate, functional, generative and measurable, thus aligning with the SMART criteria. Pretti-Frontczak & Bricker (2000) provide an overview of quality dimensions of goals set for children in special education provision and refer to functionality, generality, instructional context, measurability and hierarchical relationship as key features. These also match with the SMART definitions: functionality is present when the goal enables children to successfully negotiate her environment with as much independence as possible (matching the dimensions of *attainability* and *relevance* of the SMART acronym); generality refers to the quality of goals that while being specific to certain skills can be applicable in various contexts (matching the dimension of *specificity* and *relevance* of the SMART acronym); the instructional context refers to whether or not the goal and objective can be frequently and easily targeted across daily routines in a meaningful way (matching the *attainability* and *relevance* criteria of the SMART acronym); measurability refers to whether goals and objectives are observable (can be seen, heard, counted) and therefore measurable (matching the *measurability* criterion of SMART); and lastly, hierarchical relationship refers to whether the specific goals are related to more general goals, thus serving as precursors or building blocks (matching the criterion of *attainability* of SMART).

Despite some available evidence in the literature of goal setting for children with disabilities, there is no consensus in England on how to develop the regulated outcomes for children with EHC plans (Palikara, Castro, Gaona & Eirinaki, 2018). The Department of Education has released periodic reports on the implementation of EHC plans from 2015 (DfE, 2017; DfE, 2017b), however these reports focus mostly on numbers of EHC plans completed and on the views of a small number of parents and professionals about the process, rather than on an extensive examination of the quality of the content of the plans themselves. More recent research commissioned by the same Department of Education, although providing a very comprehensive picture of parental perspectives on the EHC planning process (Adams et al., 2017), does not encompass an objective content analysis of the final products released, and no studies to date have focused specifically on the quality of the outcomes designed. This is the first study providing evidence on the nature of the outcomes written for children with SEND in England, thus highlighting important indicators for quality assurance in the EHC planning process; on a wider level, this study provides important data regarding how the new

English special education policy and provision meets international standards, such as *participation* and *functioning*-driven service provision, as indicators for inclusion.

Thus, the purpose of this study was to provide evidence on the content and quality of the outcomes defined for children with SEND in England, as included in their EHC plans, to identify areas for improvement, and to examine the role of a number of variables on the quality of these outcomes, such the local authority where the EHC plan came from, the type of school that the child attended and the condition/diagnosis/main need of the child.

Specifically, we aimed to understand:

- a) The overall quality of the outcomes written for children with SEND in England, based on how functional and participation-based they are;
- b) Whether the content of the outcomes depends on the local authority and/or type of school that the child attends;
- c) Whether short-term outcomes differ from long-term outcomes in relation to their quality;
- d) Whether the content of the outcomes in the EHC plans varies according to the type of need that the children present.

Table 1. Overview of quality indicators of outcomes for children with special educational needs and disabilities and their match to the SMART acronym definitions

SMART criteria	McWilliam (2009) <i>functional goals in special education and early intervention</i>	Snyder, Rakap, Hemmeter, McLaughlin, Sandall, & McLean (2015) <i>Embedded instruction model</i>	Pretti-Frontczak & Bricker (2000) <i>quality dimensions of goals</i>
Specific (Siegert & Taylor, 2004; Jung, 2007)	Should specify a daily routine in which the child will participate;	Should specify who the learner is, the behaviour we are looking for, the conditions in which the behaviour should occur, the activities performed	Generality

	Should specify precisely what the child is expected to do	to achieve that behaviour or skill and	
Measurable (Conzemius & O'Neill, 2011)	There should be an acquisition criterion (when do we know that the child has acquired the behaviour?) There should be a generalization criterion (i.e. to other contexts and situations)	Should specify the criterion for acquisition	Measurability
Attainable (Piskurich, 2015)	The acquisition criterion should be relevant for the child's participation, i.e., embedded in a routine	Should specify the conditions in which the behaviour should occur	Functionality Hierarchical relationships
Relevant (Siegert & Taylor, 2004; Jung, 2007)	The acquisition criterion should be relevant for the child's participation, i.e., embedded in a routine		Instructional Context Generality
Time-framed (Conzemius & O'Neill, 2011)	There should be a timeframe for the acquisition of the new skill or behaviour		

2. Method

2.1 Sample and procedures

This study was part of a larger project looking at various aspects of the implementation of the new SEND policy in England, particularly focusing on outcomes developed for children in their EHC plans. Participant recruitment started with sending letters to London local authorities in close proximity to the area where the research team is based. Because the process of finalising EHC plans between 2015 and 2018 was slower than predicted by the government, the sampling area has expanded so that the target number of at least 200 EHC plans could be achieved. Once a meeting was agreed with the SEND representative for each local authority in order to obtain their support and agreement regarding participation in this research, meetings were arranged with individual schools. Those schools that have agreed to participate in the study, liaised with parents and guardians of children and young people with EHC plans by sending information sheets and consent forms about the research project. Opt-in consent forms were sent by the parents/guardians directly to either the research team or to the respective school, after which the EHC plan would be released in hard copy or digital copy, as per the family and school's preference. All digital copies were kept in password protected files and hard copies in locked cabinets, only accessible by the core research team. Ethical approval to undertake this study was obtained by the Ethics Committee of the hosting research institution.

The sample in this study was comprised by 236 EHC plans, which provided 2813 outcomes (see table 2 for an overview of the frequency of plans per type of need, type of placement, gender and key stage). Even though the target number for this study was 200 plans, we have received more consent forms than predicted, and therefore have included all received plans in the study (at recruitment stage the research team could not estimate how many plans would be received from each school, so the strategy adopted was to engage with as many as possible). The plans came from 42 different schools in 11 local authorities and belong to 69 girls and 167 boys from 4 to 21 years of age. There were 36 additional EHC plans that were received by the research team and excluded from the analysis for a variety of reasons: they were not readable in the format sent, they did not provide outcomes as they were still in draft mode, or they were not part of a London local authority (for consistency purposes, these were excluded). When looking at the Income Deprivation Affecting Children Index (IDACI) of the participating local authorities in this study, two of them are at the top 5% of the whole country, meaning they are some of the most deprived regions in what

concerns children’s lives and wellbeing - LA 7 and 9; LAs 1, 2 and 3 are at the top 20% in the country regarding their IDACI and LA 8 is at the bottom 10%, meaning this is one of the most affluent in the country. We recorded the main type of need that the children present, as described in their EHC plans and following the typology adopted by the UK Government on the school census 2017-2018 (DfE, 2018): 86 children have a diagnosis of Autism Spectrum Disorder (ASD), 32 have a Speech Language and Communication disorder (SLC), 8 children have Social Emotional and Mental Health difficulties (SEMH), 7 children have Specific Learning Disabilities (SLD), 2 children have Hearing Impairment (HI), 2 children have Moderate Learning Disability (MDL), 2 have Multi-Sensory Impairment (MSI), 2 have Physical Disabilities (PD) and 95 children have other difficulties, such as Global Developmental Delay, Attention Deficit Hyperactivity Disorder, Epilepsy, Genetic Syndromes or others not-specified (coded as OTHER).

In terms of age ranges, the EHC plan sample is comprised of: 52 plans referring to children in Key Stage 1 (from 4 to 7 years of age) and respective 537 outcomes; 59 children in Key stage 2 (from 8 to 11 years old) and respective 931 outcomes; 63 children in Key Stage 3 (from 12 to 14 years old) and respective 741 outcomes; 15 children in Key stage 4 (from 15 to 16 years of age) and respective 146 outcomes; and 29 children in post-16 education (from 16 to 21 years of age) and respective 303 outcomes. For 18 children it was not possible to access their age, given the lack of consent to do so.

Table 2. Frequency of plans per type of need, type of placement, gender and key stage

Sample characteristics		Frequency of children/plans
Type of need	Autism Spectrum Disorder (ASD) Speech Language and and	86
	Communication disorder (SLC)	32
	Social Emotional and Mental Health difficulties (SEMH)	8
	Specific Learning Disabilities (SLD)	7

	Hearing Impairment (HI)	2
	Moderate Learning Disability (MDL)	2
	Multi-Sensory Impairment (MSI)	2
	Physical Disabilities (PD)	2
	Other such as Global Developmental Delay, Attention Deficit Hyperactivity Disorder, Epilepsy, Genetic Syndromes or others not-specified (coded as OTHER).	95
Placement	Mainstream	125
	Special setting	111
Gender	Male	167
	Female	69
Key stage	KS1	52
	KS2	59
	KS3	63
	KS4	15
	Post-16	29
	unknown	18

2.2 Measures and data analysis

All outcomes from the 236 EHC plans gathered were inputted on a database and independently rated by two experienced and trained researchers. For the rating, the Goal Functionality Scale III, based on the model by McWilliam (2009), was used. This model was selected for being the most extensive framework of quality indicators for outcomes available with a matching scale for the assessment of outcomes. This scale was developed to evaluate the quality of outcomes written for children with Special Educational Needs, and has been used in similar studies with very high levels of inter-rater agreement (e.g. Boavida, Aguiar, McWilliam and Serpa Pimentel, 2010); The scale is comprised of 7 items: (a) indication of

the routine in which the child will participate [criterion 1], (b) specification of the desired behaviour [criterion 2], (c) relevance of the specified behaviour for the child's overall participation [criterion 3], (d) quantification of the acquisition criterion [criterion 4], (e) relevance of the acquisition criterion (included in a daily routine) [criterion 5], (f) presence of a generalization criterion [criterion 6], and (g) presence of a timeframe criterion [criterion 7]. Each outcome is rated independently on a scale of 1 to 4: *not at all*, *somewhat*, *much*, or *very much*. 10% of the outcomes were randomly selected using an automatic number generator and cross-checked by the two coders; where agreement was not reached, a third judge with similar expertise was called to support decision-making regarding the final coding. Table 3 presents examples of outcomes and their respective ratings by the two independent coders.

Descriptive frequencies were ran for all variables under analysis. Chi-square tests were performed to examine the independence of distributions between the categorical variables being studied. Bootstrap was performed due to small cell frequency in some cases, which is considered best practice, as even for small-cell frequencies these tests are considered to maintain the nominal level quite accurately (Jyh-Jiuan Lin, Ching-Hui Chang & Nabendu Pal, 2015).

Table 3. Examples of outcomes and their agreed ratings against the criteria set by the Goal Functioning Scale II (McWilliam, 2009)

Criteria according to McWilliam (2009)	Outcome 3, EHC plan 1 <i>K will follow and retain sentences with four key words within so that he is able to follow basic instructions and engage more successfully in aspects such as stories and following classroom instructions</i>	Outcome 3, EHC plan 2 <i>C will develop his language skills to a level where he can use subject-verb combinations to comment on events in structured activities</i>	Outcome 4, EHC plan 214 <i>The Pupil to recognise familiar words or symbols up to 100</i>	Outcome 19, EHC plan 214 <i>Reducing the Pupil's inappropriate grabbing of members of the public or strangers when in public spaces.</i>
(a) indication of the routine in which the child will participate [criterion 1]	<i>Much</i>	<i>Somewhat</i>	<i>Not at all</i>	<i>Much</i>
(b) specification of the desired behaviour [criterion 2]	<i>Somewhat</i>	<i>Somewhat</i>	<i>Somewhat</i>	<i>Much</i>
(c) relevance of the specified behaviour for the child's overall participation [criterion 3]	<i>Very much</i>	<i>Not at all</i>	<i>Much</i>	<i>Very Much</i>
(d) quantification of the acquisition criterion [criterion 14]	<i>Somewhat</i>	<i>Not at all</i>	<i>Not at all</i>	<i>Not at all</i>

(e) relevance of the acquisition criterion (included in a daily routine) [criterion 5]	<i>Not at all</i>	<i>Not at all</i>	<i>Not at all</i>	<i>Not at all</i>
(f) presence of a generalization criterion [criterion 6]	<i>Not at all</i>	<i>Not at all</i>	<i>Not at all</i>	<i>Not at all</i>
(g) presence of a timeframe criterion [criterion 7]	<i>Not at all</i>	<i>Not at all</i>	<i>Not at all</i>	<i>Not at all</i>

3. Results

This study aimed to provide a large-scale evaluation of the quality of the outcomes defined for children who are in receipt of an Education, Health and Care plan in Greater London, England, having done so by examining 236 EHC plans. Inter-rater agreement was 77.1%. However, if we consider differences between two categories of rating – the ones of low quality which were rated as *somewhat* and *not at all*, and the ones of higher quality rated as *much* and *very much*, agreement is 96.5%. There were only 7 outcomes in which the two raters disagreed on the overall quality – high or low. Average intraclass correlation (ICC2) agreement between the raters from the two judges, across the 7 criteria of the GFS was .49, which is considered *fair* (Chicchetti, 1994). However once again, if we look at what has been considered a high-quality outcome (*much* and *very much*) and a low quality outcome (*not at all* and *somewhat*), ICC2 was .65, which is considered *good* (Chicchetti, 1994).

What is the overall quality of the outcomes written for children with SEND in England?

Table 4 presents the frequency of ratings for all 2813 outcomes based on the GFS criteria. Across all seven items of the scale, the majority of the ratings fell onto the ‘*not at all*’ and ‘*somewhat*’ categories, suggesting that the quality of the outcomes is considerably low. This is particularly apparent when looking at whether the outcomes include a generalisation criterion, which was not incorporated in any of the analysed outcomes. The criteria that received more frequent ratings of ‘*much*’ and ‘*very much*’ were: *emphasizes the child’s participation in a routine* [criterion 1], *states specifically what the child will do* [criterion 2] and *addresses a skill that is either necessary or useful for participation in home, school, or community routines* [criterion 3]; however, even in these items, the great majority of the ratings fell on ‘*not at all*’.

Table 4. Frequency of ratings for all 2813 analysed outcomes based on the Goal Functionality Scale III (McWilliam, 2009)

	Criterion 1	Criterion 2	Criterion 3	Criterion 4	Criterion 5	Criterion 6	Criterion 7
	Emphasizes the child's participation in a routine (i.e., activity)?	States specifically (i.e., in an observable and measurable manner) what the child will do?	Addresses a skill that is either necessary or useful for participation in home, "school," or community routines?	States an acquisition criterion (i.e., an indicator of when the child can do the skill)?	Has a meaningful acquisition criterion	Has a generalization criterion	Has a criterion for the timeframe
<i>Ratings</i>							
<i>not at all</i>	2295	2163	2108	2598	2768	2813	2529
<i>somewhat</i>	449	506	547	175	40	0	284
<i>much</i>	62	133	142	36	4	0	0
<i>very much</i>	7	11	16	4	1	0	0
<i>total</i>	2813	2813	2813	2813	2813	2813	2813

Does the overall quality of the outcomes depend on the local authority and/or type of school that the child attends?

Some aspects of the quality of the outcomes vary between local authorities, with some of them developing higher quality outcomes than others in relation to almost all criteria, with most effect sizes ranging from medium to large¹, and one small. This was observable in relation to *specifying a routine in which the child will participate* [criterion 1] ($\chi^2(30) = 413.42, p < .001, V = .22$), *specifying what the child should be doing* [criterion 2] ($\chi^2(30) = 446.73, p < .001, V = .23$), *in targeting a skill that is relevant for participation* [criterion 3] ($\chi^2(30) = 508.74, p < .001, V = .25$), *in mentioning a specific acquisition criterion* [criterion 4] ($\chi^2(30) = 229.38, p < .001, V = .17$), *in mentioning the relevance of the acquisition criterion* [criterion 5] ($\chi^2(30) = 75.05, p < .001, V = .09$) and *in mentioning a timeframe for acquisition* [criterion 7] ($\chi^2(10) = 826.74, p < .001, V = .55$). There were no differences observed between local authorities regarding the *mention to a generalisation criterion* [criterion 6], since this was not included in any of the analysed outcomes. It should be noted

¹ Effect size is given by Cramer's V, interpreted as .1 = small effect; .3 = medium effect; .5 = large effect (Murphy & Myers, 1998).

that when considering the IDACI ratings of the Las involved in the study, the bottom ranked Las (and therefore most affluent) are the ones with outcomes rated as having higher quality.

Some aspects of the quality of the outcomes vary depending on the type of school where the plan comes from, with effect sizes ranging from small to medium; In terms of *specifying a routine in which the child will participate* [criterion 1] ($\chi^2 = 39.01, p < .001, V = .12$), mainstream schools have 97.8% of their outcomes being rated with 'not at all' and 'somewhat', while special schools seem to have a wider range of ratings, with 85% of their outcomes rated as 'not at all' and 'somewhat' and 15% rated as 'much' and 'very much'; this wider distribution of ratings in special schools was also observed in relation to *specifying the behaviour or skills that the child should be presenting* [criterion 2] ($\chi^2 (3) = 77.97, p < .001, V = .17$); in *targeting a skill that is relevant for participation* [criterion 3] ($\chi^2 (3) = 99.05, p < .001, V = .19$); in *mentioning a specific acquisition criterion* [criterion 4] ($\chi^2 (3) = 15.58, p < .001, V = .08$); and in *mentioning a timeframe for acquisition* [criterion 7] ($\chi^2 (3) = 11.96, p < .001, V = .07$). There were no differences observed between local authorities regarding the *mention to a generalisation criterion* [criterion 6], since this was not included in any of the analysed outcomes. Differences were also not found in relation to *mentioning the relevance of the acquisition criterion* [criterion 5].

Are short-term outcomes different from long-term outcomes in relation to their quality?

Only 1345 (48%) outcomes were specified as long-term or short-term; the remaining 1468 (52%) did not present any form of timeframe for acquisition.

The quality of the outcomes seems to vary depending on whether these are long-term or short-term or non-specified, where plans that only have long-term outcomes they also have lower frequency of outcomes than those that include short-term ones (average number of outcomes per plan $M=11.87, SD=7.46$). In criterion 1, which relates to *specifying a routine in which the child will participate*, non-specified outcomes seem to score lower ($\chi^2 (6) = 233.19, p < .001, V = .29$); In criterion 2 - *specifying the behaviour or skills that the child should be presenting* ($\chi^2 (6) = 254.66, p < .001, V = .30$), criterion 3 - *targeting a skill that is relevant for participation* ($\chi^2 (6) = 289.09, p < .001, V = .32$), criterion 4 - *mentioning a specific acquisition criterion* ($\chi^2 (6) = 65.47, p < .001, V = .11$) and criterion 5 - *mentioning the relevance of the acquisition criterion* ($\chi^2 (6) = 13.05, p < .001, V = .05$), long-term outcomes are of lower quality than non-specified and short-term outcomes; regarding

criterion 7 (*timeframe*) short-term outcomes are of better quality than non-specified or long-term ($\chi^2(2) = 130.75, p < .00, V = .22$), as it would be expected. Here, it is important to clarify that criterion 7 of the GFS refers to whether a specific timeframe has been explicitly included in the formulation of the outcome, which is exclusive of that behaviour, regardless of whether the outcome has been classed as 'long-term' or 'short-term' by those involved in the writing process. There were no differences regarding criterion 6 – the *mention to a generalisation criterion*.

Does the overall quality of the outcomes in the EHC plans vary according to the main type of need that the children present?

The distribution of the ratings regarding the quality of the outcomes varies, to some extent, depending on the type of need that the children present, albeit with small effect sizes. In what concerns the *specification of the routine that the child will participate in* [criterion 1], outcomes written for children with multi-sensory impairment and children with speech, language and communication difficulties have a wider distribution and therefore are more often rated as having higher quality, than outcomes written for children with other types of need ($\chi^2(9) = 55.68, p < .00, V = .11$). In relation to *specifying the behaviour or skills that the child should be presenting* [criterion 2], outcomes written for children with multi-sensory impairment and outcomes written for children with severe learning difficulties have a wider distribution and were more often rated as having more quality ($\chi^2(9) = 61.31, p < .00, V = .12$). When looking at whether the outcomes were *targeting a skill that is relevant for participation* [criterion 3], those written from children with multi-sensory impairment, social emotional and mental health needs, and specific learning difficulties were more often rated as having higher quality ($\chi^2(9) = 51.25; p < .00, V = .11$). In relation to the remaining criteria no significant differences were observed between the distributions of outcomes depending on the type of needs, as ratings were low across groups. Outcomes written for children in other groups such as autism spectrum disorders and physical disabilities were consistently rated as having lower quality, across criteria.

4. Discussion

The purpose of this study was to provide evidence on the quality of the outcomes defined for children with SEND who have an EHC plan, in England. Results have shown markedly low quality outcomes have been developed for these children, raising concerns

about the quality of provision under the new SEND policy. However, results have also shown that the quality of the outcomes depends on the local authority where they were developed, the type of school that the child attends, whether the outcomes are specified as short-term or long term and, to some extent, on the type of need of the children.

Looking at the local authorities that developed these plans, we were able to observe that those that have lower IDACI positions and are, therefore, more affluent areas, seem to perform slightly better in some quality indicators, than higher IDACI ranking local authorities; It is still yet to be understood though, the reasons for this phenomenon: it might be that more investment in training and continuing professional development for the professionals involved in the EHC process results in better practices in the most affluent local authorities; or it might be that more educated parents in these areas engage more with the EHC process, therefore providing an input on the final version of the plan that parents in less affluent areas may not have. Additionally, it would be important to investigate if more affluent local authorities employ more experienced professionals or professionals with more expertise in SEND, which could help to justify the slightly higher quality observed. Lastly, it would perhaps be important for future research to look at IDACI information at postcode level, where this is available, for a more fine-grained examination of affluent *versus* less affluent areas. These are questions that this study raises and that we recommend should be followed up in future research. An interesting avenue could also be the comparison of draft EHC plans with their final versions, following parental and child input, in order to investigate the contribution of these to the final product.

When looking at the type of school and how this can influence the distribution of the quality ratings of the outcomes, it was found that even though both types of schools (mainstream and special schools) have overall low-quality ratings, special schools have a higher frequency of high scores in some criteria when compared to mainstream schools – special schools seem to be better at *specifying a routine in which the child will participate* [criterion 1], at *specifying the behaviour or skills that the child should be presenting* [criterion 2], at *targeting a skill that is relevant for participation* [criterion 3], at *mentioning a specific acquisition criterion* [criterion 4] and at *mentioning a timeframe for acquisition* [criterion 7]. A potential explanation for this is the presence of higher-level professional expertise in special schools. This raises concerns around training for professionals working towards inclusion in mainstream education, which has been on the international agenda for at least the past three decades, since the Salamanca statement (UNESCO, 1994). However, it has been shown that although inclusion remains an ideal, it is often categorised as utopian

thinking, with education officers and headteachers expressing their concern regarding the capacity of mainstream education to meet the needs of children with various levels of severity (Croll & Moses, 2000). Local policy makers should give further consideration to the design of guiding frameworks for professional development, based on scientific research, especially tailored for those who work with children with SEND in mainstream contexts. The embedded instruction model, for instance, provides not only a framework for the design of quality outcomes, but also an embedded professional development programme based on coaching approaches, which have been proved most successful in changing practices in the classroom (Barton, Bishop & Snyder, 2014; Bishop, Snyder & Crow, 2015). A practice-based approach where a coach provides embedded and continuous feedback and feedforward (e.g. Snyder, Hemmeter & Fox, 2015) might constitute an invaluable asset in a context where inclusion is key and inter-professional collaboration must exist. Kauffman and Badar have also expressed the need to recognise different needs in order to be able to meet them and promote inclusion for all (Kauffman & Badar, 2014). This issue is of particular importance if we consider the UK results in the 2013 Teaching and Learning International Survey (TALIS; Organisation for Economic Co-operation and Development, 2013), where UK teachers seem to be working many more hours than most OECD countries, but doing so in isolation, ranking at the bottom of the scale for collaboration with others. Recent research supports this finding, with professionals reporting that they feel they lack the specialised training to effectively implement the new policy and provide holistic education, health and social care provision for all (Palikara, Castro, Gaona & Eirinaki, 2018). This result supports the idea that inclusion policy should be revisited in England, with further consideration for the integration of professional development programmes which are embedded and standardised across geographical areas.

In this study, the quality of the outcomes also varied across criteria 1, 2 and 3 depending on the children's main type of need. It is interesting to observe that outcomes for children with multi-sensory impairments have been more often rated of higher quality across all three criteria, while outcomes written for children with autism spectrum disorders, for example, have very narrow distributions (and quite low ratings). One possible explanation for this fact is that it may be easier to write SMART outcomes and specify the routines where these will be targeted when working with children that might have higher functioning profiles; here, emergent skills may be more observable and short-term achievements more likely to be seen when targeting these skills (note that the criteria where differences were observed relate to the actual behaviour being observed and respective routine when the

behaviour will become apparent). Children that have been diagnosed with autism spectrum disorders will often have more severe or lower functioning behavioural patterns, where emergent skills may not be so obvious for inclusion in achievable outcomes. However, this would contradict findings of previous studies in the field: Boavida, Aguiar and McWilliam (2014) found that outcomes written for children with more severe disabilities seem to score higher on measurability. This phenomenon requires further investigation with the specific aim of examining the role of the type of need on the quality of the plans, with more homogenous samples across types of needs than those obtained here. However, the key to understanding how quality ratings may be dependent on the children's main type of need would be to have a label based on functioning to designate the main need, rather than on diagnosis; diagnostic labels may be masking the real functioning level of the child, as it is well recognised that children with the same diagnosis can have very different functioning profiles (Castro & Pinto, 2015; Pinto et al., 2018). It would be important to keep a record of the child's level of functioning and participation pattern, rather the diagnostic category, and further analysis on the quality of the plans could then be conducted based on this.

4.1. Limitations and future research

Some limitations of this study should be mentioned. Even though this is a large sample of 236 plans across various local authorities which are very different from each other, the study was based in Greater London, and therefore, some of the results observed might not correspond to the situation across the whole country. However, previous research at national level has highlighted that these concerns are shared by professionals from all over the country (Palikara, Castro, Gaona & Eirinaki, 2018). Future research should look at EHC plans conducted in other regions of the country to confirm this trend and/or use more fine indicators of wealth across local authorities. Additionally, it would be of interest to compare national data with data from other countries with their own policies and supports services, with differences analysed in light of those discrepancies and similarities in global approaches to SEND provision.

The GFS II (McWilliam, 2009) was chosen as the tool to guide the evaluation of outcomes given its exhaustive framework which encompasses several quality dimensions, overlapping the SMART criteria, however, more research into its psychometric properties is needed. Additionally, we argue that future research should consider the exploration and design of a tool to evaluate outcomes specifically formulated to be included in EHC plans.

Such tool would support the standardisation of provision and help raise its quality across contexts.

One important limitation relates to the effect sizes reported which, at times, are quite small, in particular regarding to differences that depend on the child's main type of need. Therefore, interpretations about these differences should be made with care. We argue that future research should perhaps look at differences in quality of plans and outcomes between children with different participation patterns and/or functioning profiles, instead of children with different diagnosis.

Lastly, one clarification should be made: the current study evaluated the extent to which the outcomes included in the EHC plans gathered were functional according to the criteria established by McWilliam (2009); Even though we have argued that there is a match between the criteria proposed by McWilliam (2009) and the SMART acronym, we did not evaluate the extent to which the outcomes were SMART; the reason for this is two-fold: first, because there is no specific standardised tool for this effect, and secondly, because the extent of *SMARTness* of an outcome might not necessarily reflect how this is going to be implemented. The results of this study are important as they provide evidence regarding the need for improving the quality of outcome design for children with SEND, but further research should be undertaken regarding the impact of high-quality outcomes in the quality of actual service provision and children's participation.

5. Conclusion

In sum, three points can be made as a conclusion for this study: first, if the EHC plans should contain high-quality outcomes (potentially SMART), developed collaboratively between education health and social care and focusing on *participation*, there is a need for effective professional development and training for those involved in the process to increase specialisation and consequently the quality of these outcomes; secondly, such training should be standardised to all professionals involved in the process, across local authorities and regardless of the type of school, to ensure a fairer provision to all children with SEND, regardless of their socio-economic background or of the severity of their needs. Third, if such training is to be implemented, it should be based on international evidence available around the most effective models for professional development; for instance, research has shown that coaching models of professional development for staff working in SEND favour the development of professional skills such as designing SMARTer outcomes (Barton, Bishop & Snyder, 2014). Lastly, it is crucial that the quality issues highlighted around outcomes in the

English EHC plans are not neglected by policy makers; the English special education system has been highly regarded for decades, but the current national picture shows a lack of compliance with global principles of inclusion.

Acknowledgements

We are thankful to all the families and children who agreed to release their EHC plans to be included in the analysis. We thank all professionals who have facilitated data collection.

Funding:

This work was supported by the British Academy/Leverhulme Trust small grant scheme [grant reference SG142214];

References

- Adams, L., Tindle, A., Basran, S., Dobie, S., Thomson, D., IFF Research, Robinson, D. & Shepherd, C. (2017). Experiences of Education, Health and Care plans: a survey of parents and young people. University of Derby for the Department of Education.
- Adolfsson, M. (2013). Applying the ICF-CY to identify children's everyday life situations: A step towards participation-focused code sets. *International Journal of Social Welfare*, 22(2), 195-206.
- Barton, E., Bishop, C., & Snyder, P. (2014). Quality instruction through complete learning trials: Blending intentional teaching with embedded instruction. In K. Pretti-Frontczak, J. Grisham-Brown, & L. Sullivan (Eds.), *Blended practices for all children*. *Young Exceptional Children Monograph*, 16, 73-96.
- Bishop., Snyder, P., & Crow, R. (2015). Impact of video self-monitoring with graduated training on implementation of embedded instructional learning trials. *Topics in Early Childhood Special Education*, Online First, doi: 10.1177/0271121415594797
- Boavida, T., Aguiar, C., & McWilliam, R. A. (2014). A training program to improve IFSP. *IEP goals and RA McWilliam*.

- Boavida, T., Aguiar, C., McWilliam, R. A., & Pimentel, J. S. (2010). Quality of individualized education program goals of preschoolers with disabilities. *Infants & Young Children, 23*(3), 233-243.
- Boesley, L., & Crane, L. (2018). 'Forget the Health and Care and just call them Education Plans': SENCO s' perspectives on Education, Health and Care plans. *Journal of Research in Special Educational Needs*.
- Bornman, J. (2017). Human rights of children with disability. An Emerging Approach for Education and Care: Implementing a Worldwide Classification of Functioning and Disability.
- Castro, S., & Palikara, O. (2016, November). Mind the gap: the new special educational needs and disability legislation in England. In *Frontiers in Education* (Vol. 1, p. 4). Frontiers.
- Castro, S., & Pinto, A. (2015). Matrix for assessment of activities and participation: measuring functioning beyond diagnosis in young children with disabilities. *Developmental neurorehabilitation, 18*(3), 177-189.
- Cicchetti, Domenic V. (1994). "Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology". *Psychological Assessment, 6* (4): 284–290. doi:10.1037/1040-3590.6.4.284.
- Croll, P., & Moses, D. (2000). Ideologies and utopias: Education professionals' views of inclusion. *European Journal of Special Needs Education, 15*(1), 1-12.
- Department for Education and Department of Health (2015). SEND Code of Practice 0–25 Years. *Statutory Guidance for Organisations which Work with and Support Children and Young People with Special Educational Needs and Disabilities*.
- Department for Education (2017a). Education, health and care plans: parents and young people survey.
- Department for Education (2017b). Statements of SEN and EHC plans: England, 2017.
- Department for Education (2018). School census 2017 to 2018 Guide, version 1.8.
- Department for Education and Department of Health and Social Care. (2014). SEND code of practice: 0 to 25 years.
- European Agency for Development in Special Needs Education [Internet]. Assessment resource guide. 2010. Available from: <http://www.europeanagency.org/agencyprojects/assessment-resource-guide> R
- Hollenweger, J. (2017). Applying ICF in education and care. In *An Emerging Approach for Education and Care* (pp. 35-48). Routledge.

- Imms, C., Adair, B., Keen, D., Ullenhag, A., Rosenbaum, P., & Granlund, M. (2016). 'Participation': a systematic review of language, definitions, and constructs used in intervention research with children with disabilities. *Developmental Medicine & Child Neurology*, 58(1), 29-38.
- Jung, L. A. (2007). Writing SMART objectives and strategies that fit the ROUTINE. *Teaching Exceptional Children*, 39(4), 54-58.
- Jyh-Jiuan Lin, Ching-Hui Chang & Nabendu Pal (2015) A Revisit to Contingency Table and Tests of Independence: Bootstrap is Preferred to Chi-Square Approximations as Well as Fisher's Exact Test, *Journal of Biopharmaceutical Statistics*, 25:3, 438-458, DOI: 10.1080/10543406.2014.920851
- Kauffman, J. M., & Badar, J. (2014). Instruction, not inclusion, should be the central issue in special education: An alternative view from the USA. *Journal of International Special Needs Education*, 17(1), 13-20.
- Legislation.gov.uk. [Internet]. Children and Families Act 2014. [cited 11 Jan. 2015]. Available from: <http://www.legislation.gov.uk/ukpga/2014/6/contents/enacted>
- McKay, E., Mahon, D., Donellan, G., Haracz, K., Sheldon, S., & Ryan, S. (2017). 17 Participation and inclusion. *Participation in Health and Welfare Services: Professional Concepts and Lived Experience*, 209.
- McWilliam, R. A. (2005). Goal Functionality Scale. Nashville, TN: Vanderbilt University Medical Center.
- Murphy, K.R., and Myers, Brett, 1998, *Statistical power analysis—A simple and general model for traditional and modern hypothesis tests*: Mahwah, N.J., Lawrence Erlbaum Associates, Inc.
- Organisation for Economic Co-operation and development (2014). *A Teachers' Guide to TALIS 2013: Teaching and Learning International Survey*. TALIS, OECD. Publishing. <http://dx.doi.org/10.1787/9789264216075-en>.
- O'Neil, J. & Conzemius, A. (2006). *The Power of SMART Goals : Using Goals to Improve Student Learning*. Solution Tree Press. p. 33. ISBN 978-1-932127-87-4.
- Palikara, O., Castro, S., Gaona, C., & Eirinaki, V. (2018). Professionals' views on the new policy for special educational needs in England: ideology versus implementation. *European Journal of Special Needs Education*, 1-15.
- Pinto, A. I., Grande, C., Coelho, V., Castro, S., Granlund, M., & Björck-Åkesson, E. (2018). Beyond diagnosis: the relevance of social interactions for participation in inclusive preschool settings. *Developmental neurorehabilitation*, 1-10.

- Pretis, M. (2017). 'Let Us Be Prepared, but Wait and See'the Use of ICF-CY in Early Childhood Intervention and Paediatric Social Care in Germany and Neighbouring Countries. An Emerging Approach for Education and Care: Implementing a Worldwide Classification of Functioning and Disability, 165-177.
- Pretti-Frontczak, K., & Bricker, D. (2000). Enhancing the quality of individualized education plan (IEP) goals and objectives. *Journal of early intervention*, 23(2), 92-105.
- Rakap, S. (2015). Quality of individualised education programme goals and objectives for preschool children with disabilities. *European Journal of Special Needs Education*, 30(2), 173-186.
- SAKAI, Y., Tanaka, K., & TOKUNAGA, A. (2017). The use of the ICF-CY in special needs education in Japan. In *An Emerging Approach for Education and Care* (pp. 83-96). Routledge.
- Siegert, J. R. & Taylor, W.J. (2004). Theoretical aspects of goal-setting and motivation in rehabilitation. *Disability & Rehabilitation. Informa*. 26 (1): 1-8.
- Simeonsson, R. J. (2009). ICF- CY: a universal tool for documentation of disability. *Journal of Policy and Practice in Intellectual Disabilities*, 6(2), 70-72.
- Simeonsson, R. J., Scarborough, A. A., & Hebbeler, K. M. (2006). ICF and ICD codes provide a standard language of disability in young children. *Journal of Clinical Epidemiology*, 59(4), 365-373.
- Simeonsson, R.J. (2006) Appendix C: Defining and Classifying Disability in Children. In Field MJ, Jette A, & Martin, L. *Workshop on Disability in America: A New Look*. Washington, DC: National Academy Press.pgs.67-87.
- Snyder, P., Hemmeter, M. L., & Fox, L. (2015). Supporting implementation of evidence-based practices through practice-based coaching. *Topics in Early Childhood Special Education*, Online First, doi:10.1177/0271121415594925
- Snyder, P. A., Rakap, S., Hemmeter, M. L., McLaughlin, T. W., Sandall, S., & McLean, M. E. (2015). Naturalistic instructional approaches in early learning: A systematic review. *Journal of Early Intervention*, 37(1), 69-97.
- Tod, J., Castle, F., & Blamires, M. (2013). *Individual Education Plans Implementing Effective Practice*. Routledge.
- Unesco. (1994). *The Salamanca Statement and Framework for action on special needs education: adopted by the World Conference on Special Needs Education; Access and Quality*. Salamanca, Spain, 7-10 June 1994. Unesco.

- VanDerHeyden, A. M., Snyder, P., Smith, A., Sevin, B., & Longwell, J. (2005). Effects of complete learning trials on child engagement. *Topics in Early Childhood Special Education*, 25(2), 81-94.
- Warnock, M. (2005). Pamphlet "Special educational needs: a new look", n. °11 in a series of policy discussions. Published by the Philosophy of Education Society of Great Britain.
- World Health Organization. (2001). *International classification of functioning, disability and health: ICF*. Geneva: World Health Organization.
- World Health Organization. (2007). *International classification of functioning, disability and health: Children and youth version: ICF-CY*. World Health Organization.