

**Symposium:
Pedagogical Approaches for Inclusion: An Evidence Base?**

**What pedagogical approaches can effectively include children with special educational needs in mainstream classrooms? The interactions of peers, teachers and support staff -
A systematic review process**

Jonathan Rix and Kathy Hall, The Open University

**Paper presented to the
International Special Education Congress
University of Strathclyde
1-4 August 2005**

**Correspondence:
Jonathan Rix
Faculty of Education & Language Studies,
The Open University
Milton Keynes MK7 6AA
Email: J.R.M.Rix@open.ac.uk**

Abstract

This paper addresses issues raised by the systematic literature review process. The authors are currently examining the literature on the pedagogy of mainstream teachers and support staff that effectively include children with special educational needs, with a view to assessing the interactions of peers, teachers and support staff. This paper sets out the methods of the systematic review, how we defined our terms and narrowed our focus. It explores the tensions that we confronted as part of this process. It explores in particular how we built on a previous review, and dealt with the criteria used to include and exclude studies and to carry out keywording. The paper concludes by highlighting some limitations of the systematic review process, and their impact on the ways in which we frame the reviews we create.

Introduction

Teachers are expected to work effectively with an ever-widening diversity of students, often feeling that they are unprepared to do so (Scruggs & Mastropieri, 1996; Farrell, 1997; Jenkinson, 1997). Part of the challenge for them is their understanding that there are many pedagogies written about and which make claims to effectiveness, without knowing the degree to which these teaching approaches have been appropriately researched and shown to be suitable practice. The complexity of the learning environment in which teachers work means making such judgements about pedagogies is particularly challenging, as there are many variables operating at any given moment. In an attempt to identify those techniques and materials that mainstream teachers can rely upon and use, it was decided to systematically review the research into pedagogies that effectively include people with special educational needs. This paper reports on the second year of this progressive and developing three year systematic review programme undertaken for the Teacher Training Agency (TTA).

Evans and Benefield (2001) see the key features of a systematic review as being:

an explicit research question to be addressed; transparency of methods used for searching for studies; exhaustive searches which look for unpublished as well as published studies; clear criteria for assessing the quality of studies (both qualitative and quantitative); clear criteria for including or excluding studies based on the scope of the review and quality assessment; joint reviewing to reduce bias; a clear statement of the findings of the review

(Evans and Benefield, 2001, p.529)

They suggest that systematic reviews differ from narrative forms in that they are more focussed, offering clarity about search strategies and the criteria and methodology used for including and assessing reviews. Systematic reviews are frequently described as providing answers to specific questions, whereas narrative reviews provide a map of large, complex areas where multiple issues are involved. This means that systematic reviews are seen as having more relevance to practical user issues and narrative reviews are of more interest to academics.

There have been a number of Systematic Reviews carried out in the last few years on behalf of the TTA and DfES, and three, prior to last year's review (Nind et al, 2004), which focused in particular on specific areas of special educational needs and inclusion. These reviews considered appropriate responses to pupil behaviour (Harden et al, 2003), the impact of paid adult support (Howes et al, 2003) and school-level approaches to facilitating the participation of all students (Dyson et al, 2002). There have been a number of other reviews in this area beyond the TTA/DfES remit too, for example, two meta-analyses by Quinn et al. (1999), who examined how social skills training interventions improve social skills or reduce problem behaviours amongst young people with emotional and behavioural disorders, and Stage and Quiroz (1997), who examined the effectiveness of school-based interventions in reducing disruptive classroom behaviour.

Within these studies there are a wide variety of strengths and weaknesses identified in relation to the systematic review process. Harden (2003) for example, talks about a bias in the kinds of journals they searched and the impact this would have on the range of theoretical models evaluated by the included studies, pointing to the bias towards English language papers in particular. Time restriction is also a common concern, particularly in relation to gaining access to unpublished materials (Nind et al, 2004). There are clearly a number of issues raised, too, when reviews rely solely upon bibliographic databases for searching (Evans et al, 2001, Dyson et al, 2002). For example, the use of databases such as ERIC and PsycINFO biases the review towards their limited range of topics and journals and tends to mean fewer books or book chapters are identified. There are also issues of database construction, so that there can be significant time lags between the completion of a study and its appearance on the database. The information provided varies greatly too, not only in terms of whether the databases provide a citation, abstract or a full text, but also in relation to the

keywords used. Keywording across the databases can vary greatly, as can the ability to search for more general or specific terms, synonyms and related terms.

There is often a concern expressed (Howes et al, 2003, Hammersley, 2001) about the tendency to identify quantitative research as opposed to qualitative. Papers are less likely to be seen as valid and reliable if there is not a clearly defined, measurable outcome. In addition, the a priori criteria can rule out papers that could indirectly have a bearing on the subject at hand. Similarly, issues of generalisability are a major concern. If the review is focused on a very specific set of criteria, then findings cannot be applied to a situation in which those criteria do not apply. The synthesis process at the end of review means findings can be generalised to some degree, but it would be against the spirit of systematic reviews to go further. This can limit the understanding of users, who may feel the need to generalise beyond the review's scope. This would seem to be a likely outcome in educational settings in which the variables that operate mean that no two situations are entirely the same.

The challenges of the systematic review for 2005

In 2004, the authors of this review were part of the team that carried out a systematic literature review of studies which investigated pedagogical approaches that effectively include people with special educational needs (Nind et al, 2004). This was carried out using the methods and tools created and supplied by the EPPI-Centre (Evidence for Policy and Practice Information), who aim to enable educational researchers in carrying out systematic reviewing and to disseminate findings to practitioners and policy makers.

In the first year of the review, we identified a wide variety of possible pedagogical approaches that could act as the focus for the in-depth review. When making our final choice we did not wish to focus on an individual category of need or studies that shared a broad aim, but wanted to identify approaches that could be used by ordinary teachers for the whole class without recourse to specialist resources. It was decided, therefore, to focus specifically upon peer group interactive approaches. In the second year we are not focusing upon a single cluster of approaches, but upon the interactions of peers, teachers and support staff across a range of identified approaches. This seemed a particularly relevant starting point, given the socially constructed nature of knowledge and the importance of interactions in mediating and scaffolding learning opportunities for all. Regardless of the pedagogical approach, the interactions that occur within it will be central to its success. It was also recognised that there

is a tendency to neglect this aspect of pedagogy in relation to effective inclusion (Skidmore, 2004).

We were also very aware of the breadth of the field in which we were working. In many ways our experience highlighted both a strength and weakness of the systematic review process. Our initial question in 2004 had revealed over 1800 papers that had some relevance to our database keywords. Examination of abstracts had brought this down to 450 papers, which had been further narrowed down to the final papers for keywording. Within the tight constraints of our initial inclusion and exclusion process we keyworded 68 papers and identified a wide variety of pedagogical approaches, many of which overlapped in the learning situation. We identified too a range of individuals and groupings who were central to the teaching act.

Identified range of pedagogical approaches and teaching types

Nature of the pedagogical approach (N=68)		Who does the teaching (N=68)	
Adaptation of instruction	40	Regular, mainstream teacher	40
Adaptation of materials	24	Special teacher and regular teacher in collaboration	24
Adaptation of assessment	13	Teachers with equal roles/responsibilities in collaboration	3
Adaptation of classroom environment	17	Learning support assistant	5
Behavioural/programmatic intervention	14	Peers	23
Computer based	2	Special Education teachers	7
Peer tutoring	17	Research/research assistant	7
Peer group interactive	24	Parent	1
Team-teaching	11	Not stated	1
Other (often mixed methods)	10		

It seemed as if we could approach this diversity in two ways. Either we could focus on another pedagogical type or teaching group and examine how they operated or we could consider an aspect of the papers that cut across all the pedagogical approaches and teacher types. It was felt that the former approach could identify and explore an important part of the overall picture of effective pedagogy, helping to build up a wider picture, but it would not deal with the overlaps between the approaches and teacher types. The strength of the review process meant that we had a clearly defined set of papers, but the approaches explored within them were not so clear cut. It was not possible to isolate all variables and identify clear cause and effect in many papers, or to write about all the papers and produce a manageable review. But neither could we cherry pick appropriate papers to produce a broader overview as would be possible in a narrative review. For this reason we decided to take a cross sectional view of

the papers, providing ourselves with the opportunity to explore across the boundaries we had created with our keywords in the first year.

Of significance to the 2005 review is the 14% of papers that should have been screened in 2004 but were not received within the timeframe of the 2004 review or were unavailable. We decided to update the original review from 2004, so as to pick up those papers that had not been received and any others that had been published or added to the databases in the meantime. We sought papers that used:

- Terms indicating that the study is about children with special educational needs;
- Terms indicating that the study is about inclusion;
- Terms indicating that the study is about pedagogical approaches;
- Terms indicating that the study involves pupils aged between 7 and 14.

The outcome of this re-examination of the databases raised a number of issues identified in previous reviews. At the outset we had not expected that many new papers would be uncovered, but it transpired that an additional 1229 papers needed our consideration. The main bulk of these were papers from 1994-2004 that had only just been added to the databases, in particular ERIC. Like others before us we also recognised that the information provided varied greatly too, not only in terms of whether the databases provided a citation, abstract or a full text, but also in relation to the keywords used.

These 1229 papers were now screened, using the same set of inclusion and exclusion criteria as had been used the year before. These covered the scope of papers listed above, as well study types and time and language of publication:

Inclusion and Exclusion Criteria for 2004 and 2005 systematic reviews

	Exclusion criteria	Inclusion Criteria
	Not focused on pupils who experience special educational needs of some kind	Include a focus on students who experience special educational needs of some kind
	Not conducted in mainstream classrooms	Be conducted in mainstream classrooms
	Not concerned with pedagogical approaches	Include pedagogical approaches
	Not indicating pupil outcomes	Include an indication of student outcomes
	Not concerned with all or part of the 7–14 age range	Be concerned with the 7-14 age range or some part of it
	Descriptions, development of methodology or reviews other than systematic reviews	Be empirical – exploration of relationships, evaluations or systematic reviews
	Not written in English	Be written in English
	Not produced or published in or after 1994	Be published after 1994

Clearly, these criteria are likely to exclude many papers that could have relevance to teachers. For example, Norwich and Lewis (2000) carried out an explicit and transparent mapping exercise to explore the question of whether differences between SEN learner groups could be identified and systematically linked with learners' needs for differential teaching, concluding that there was a "lack of evidence to support SEN specific pedagogies" (Norwich et al, 2000, p.3). By seeking out only those pedagogies labelled as tested on people with special educational needs we could be ignoring very effective, everyday, mainstream practices. Criterion 1, therefore, could be seen as working against the notion that inclusive education is education for all, by excluding those papers that are general in their focus. Criterion 5 could also be problematic. In 2004 we realised that our in-depth study had greater reliability in relation to Primary age children. This could be as a result of excluding those papers that have a focus on 15-18 year olds. Certainly Criterion 7 is about our weaknesses as linguists, and denies us access to an unknown number of papers and original practices and perspectives. These issues raise a dilemma for systematic reviewing. A main strength of systematic reviews is their explicit procedures. It is suggested that such openness makes for high reliability of conclusions and for the process to be assessed by readers (Slavin, 1986). One has to ask, however, whether being open and honest about such shortcomings overcomes the impact of the shortcomings themselves, which result from requiring a narrow question and very specific criteria.

The questions we faced when applying the criteria

Our first challenge in applying these criteria was the quality of the abstracts and titles supplied. There was a great deal of variation in the amount and accuracy of information provided. A ground rule therefore had to be to include any paper that could not clearly be excluded because the synopsis or title was limited in use.

The inclusion and exclusion criteria created a number of discussion points when screening the papers, both when considering the abstracts and the papers themselves. As Dyson et al (2002, p.8) pointed out, "certain terms are particularly difficult to define in a way that can be operationalised for research purposes and for which reliable measures or indicators can be devised".

Criterion 2 (Be conducted in mainstream classrooms) produced three main questions: - firstly, "How should we deal with papers that were partially conducted in segregated settings?". We decided that such a paper could only be included if the intervention began in

one setting and was then carried on within the mainstream. If the strategy was enacted entirely in a segregated situation (such as a resource room or a special school) and then measured in the mainstream classroom we excluded the paper. It seemed important that a paper needed to explore the enactment in the context that mainstream staff would be working within. Our second question was “Does our definition of a classroom include other school spaces, such as the playing field in PE or the library?”. We decided that the definition of classroom was subject sensitive. If the space used in the paper was integral to the regular teaching of that lesson then it would be seen as being “within the mainstream classroom”. So, the library is a classroom if all pupils are using it regularly in relation to a particular area of learning, and a sports hall or playing field is a classroom when linked to Physical Education lessons. The third question was whether “conducted in” has to be taken literally. If a group of teachers are interviewed outside the classroom in relation to their practice in the classroom, is this research “conducted in” the classroom or outside? We felt that we would not exclude such studies as they could, if empirically supported, provide a variety of insights into ways of working. Such studies would tend therefore to be excluded under Criterion 4 or 6 if appropriate.

Criterion 3 (Be concerned with pedagogical approaches) produced three main questions; the first two being about the scheduling of classes and about whether assessment can be seen as pedagogy. We wished to take a broad view of pedagogy, defining it as including “classroom practices, personnel deployment, organisation, use of resources, classroom environment and curriculum, that is, what occurs in classrooms that can be seen to impact on participation and learning.”, but we felt that papers focusing on timetabling structures were dealing with school-wide issues that fell outside the pedagogy remit. We felt too that papers that merely focussed on ways of improving performance on normative assessments were not suitable for inclusion. Papers that dealt with assessment alone, were only included if they involved a formative component, in which the assessment was seen to feedback into the pupils learning and not merely measure an aspect of that learning. The third question we had about pedagogy concerned the use of specific text books or programmes. We decided that we would include books or programmes if there was a clear link to classroom practice, and not merely a measurement of outcome.

Criterion 4 (Include an indication of student outcomes) threw up a major challenge, particularly in relation to Criterion 6 (Be empirical – exploration of relationships, evaluations or systematic reviews); how do you include qualitatively measured outcomes that are

empirically rigorous enough? Can a survey of teacher attitudes ever match a closely observed lesson interaction? At what point does a survey of attitudes cease to be empirically sound and become anecdotal ... and at what point does an anecdotal view become a valid case study? When applying any criterion we are bound to make personal judgements about such matters,

even the most simple rule-following involves some judgement; and rational decision-making will often take the form of the interpretation of principles rather than the 'application' of rules

(Hammersley, 2001, p.546)

We felt this tension between interpreting principles and applying rules most keenly in relation to this question about the rigour of studies. Answering it was key to a central notion in our review question too – the notion of effectiveness. How could we say that a pedagogy effectively includes without a clear understanding of the outcomes and how they are being assessed? In our original protocol we decided to judge papers on the basis of their own criteria and how explicit they were. This allowed us to be fairly liberal in allowing papers past Criterion 4, but was not so helpful with Criterion 6. Criterion 6 included the subheading “exploration of relationships, evaluations or systematic reviews” however. It was the notion of relationships, that allowed us to include a variety of papers that were qualitative case-studies. We would include any papers that examined the relationship between components of the learning context, as long as they were not merely descriptions. We then had enough information to evaluate the rigour of their outcome assessment against their internal rationale. We had a much smaller dilemma in relation to criterion 4 as well. If a paper is about student-teachers on a training programme, and observes student-teachers carrying a particular pedagogy in a mainstream setting with pupils aged between 7-14, and then assesses how well those student teachers did...should we include it? Literally, speaking we don't have a criterion to exclude it. We therefore interpreted “Include an indication of student outcomes” as referring to the students in the 7-14 age range.

Criterion 5 (Be concerned with the 7-14 age range or some part of it) raised a number of small issues that were surprisingly time-consuming. Many papers do not clearly define the ages of the students involved. There is a tendency, particularly in papers from the United States, to rely upon curricular grouping to define the subjects of the research, for example, Grade 2 or Key Stage 3. When identifying the relevance of papers at the extremes of our age range this is particularly problematic. Will Kindergarten and Grade 1 include any seven year

olds or will Grade 10 include some 14 year olds? They might do. We can't tell. Equally difficult is deciding whether enough people within the study fell within the age-range to merit inclusion. If there are not ages mentioned, and the groups are at either end of the age-range this can be particularly frustrating. In coming to a decision about these cases we bore in mind the need to make statements that have valid and reliable applications to our age-range. If we could identify a significant number of pupils in the context of the research structure then we would include on this Criterion, if however the number was too small to make reasonable conclusions, or if we could not tell what the figures were, then the paper was excluded.

Preparing for the Keywording - Lessons from 2004

At the time of writing the review team have just begun the process of Keywording over 100 studies identified both last year and this. Prior to this stage, the team felt it was necessary to clarify the definitions of the Review Specific Keywords. In the previous year we had felt that we had a common view of terms such as:

- Raise academic attainment
- Enhance social interaction/involvement
- Improve behaviour

Yet when the keywording was analysed it became clear that there were a number of differences in interpretation that meant papers had to be revisited and keywords reassessed. We therefore had a full team debate about the operationalisation of the keywording. At the same time at the request of EPPI we revisited our notion of effectiveness, clarifying our view that we should assess this against the papers own criteria, rather than an imposed range. We felt that effectiveness is context specific and therefore to assess the effectiveness of any pedagogical approach we need to assess the degree to which the context is clearly defined within a paper as well as any measures of success. This context assessment would be part of the in-depth stage of the review process.

The ability to draw upon last year's experiences has strengthened the protocol for this year. It has also demonstrated the need for the review team and its external advisors to maintain regular and open communication about a wide variety of issues. This, of course, is not always that easy when there are nine people on the review team, a link person at the EPPI Centre, and ten external advisors. Collecting, filing, copying and distributing papers is a time

consuming and expensive process. Such organisational niceties have also taught us lessons. Last year, for example, we did not collate those papers that arrived after the keywording deadline. At the time we had not considered using the same primary question for this year's review. This is an oversight we will try to avoid this year.

Conclusion

Systematic reviews are currently playing a significant part in the creation of policy and practice throughout the UK. Our experiences with this process suggest that it has a useful role to play in answering contained and focused questions, but that this does not mean that one review alone can effectively serve the needs of users. Like Dyson et al (2002) we would suggest that this methodology provides one set of reviewing tools and processes amongst many.

Our concerns about the process have arisen from what in other senses can be seen as strengths:

- By defining categories a priori you rule out the unexpected, which may represent the new way forward.
- By suggesting that a priori criteria minimise the making of subjective judgements we imply that the review is objective. Throughout the process, from defining the question to creating and applying the inclusion and exclusion criteria we are relying upon our limited skills base and formalising our own biases and beliefs.
- By being open you do not overcome the shortcomings you identify.
- By focusing on what is contained within the papers you have identified you can only apply your understanding to the contexts explored in those papers.
- By focusing on what is contained within papers identified by a priori criteria you may be less likely to identify those research questions that are not being asked and research methods that are not being used.
- By focusing on empirical evidence it is likely that qualitative, theoretical and critical papers will be downplayed.
- By using systematic reviews as the benchmark for evidence-based practice we imply that everyday practice is not evidence-based.

- By providing a very tight, context specific set of findings, users may feel a need to extrapolate the findings to different contexts.

These concerns do not mean we are rejecting the value of systematic reviews. They are intended to underline the need for us to engage with these issues in the creation of our reviews. Our decision to revisit last year's papers from a different perspective is one possible way of doing this. This will allow us to start building up a cross-sectional view of the field. Similarly, other review teams can use the criteria of this review to inform them of the gaps in the wider systematic review framework. In this way a mosaic picture of effective pedagogies can be created. It would also seem sensible to include information in our appendices, as Howes et al (2003) did, about studies that were felt to be relevant but did not reach the high level of empirical reliability required.

In our attempts to provide teachers with a guide through the diverse and dispersed literature, our systematic review provides a tightly defined lens. It is harder for it, however, to provide a broader overview, and to encourage us to look beyond our expectations to find new ways of breaking down potential barriers to learning. For those of us involved in the review process, therefore, we need to engage in wide ranging discussions during the creation of the review, so as provide as many opportunities to confront individual assumptions before they become formalised in the systematic review structure.

References

- Dyson A, Howes A, Roberts B, (2002). A Systematic Review of The Effectiveness of School-Level Actions for Promoting Participation by All Students (Eppi-Centre Review, Version 1.1*). In: Research Evidence in Education Library. London: Eppi-Centre, Social Science Research Unit, Institute of Education.
- Evans, J. And Benefield, P. (2001) Systematic Reviews Of Educational Research: Does The Medical Model Fit? *British Educational Research Journal*, 27, 5, 527-541.
- Farrell, P. (1997) *Teaching Pupils with Learning Difficulties*, London: Cassell.
- Hammersley, M. (2001) On 'Systematic' Reviews of Research Literatures: A 'Narrative' Response to Evans & Benefield, *British Educational Research Journal*, 2, 5, 543-554.
- Howes, A., Farrell, P., Kaplan, I. & Moss, S. (2003) The Impact of Paid Adult Support on the Participation and Learning Of Pupils In Mainstream Schools. In: *Research Evidence in*

- Education Library. London: Eppi-Centre, Social Science Research Unit, Institute of Education.
- Jenkinson, J. (1997) *Mainstream or Special? Educating Students with Disabilities*, London: Routledge.
- Lewis, A., & Norwich, B. (2000) *Mapping A Pedagogy for Learning Difficulties*, Report Submitted To The British Educational Research Association (Bera), February 2000, As Part Of The National Events Programme 1999
- Nind, M. & Wearmouth, J. with Collins, J., Hall, K., Rix, J., Sheehy, K. (2004) *A Systematic Review of Pedagogical Approaches that can Effectively Include Children with Special Educational Needs in Mainstream Classrooms With A Particular Focus on Peer Group Interactive Approaches*. In: *Research Evidence In Education Library*. London: Eppi-Centre, Social Science Research Unit, Institute of Education.
http://eppi.ioe.ac.uk/EPPIWeb/home.aspx?page=/reel/review_groups/TTA_SEN/review1.htm
- Quinn, M., Kavale, K., Mathur, S., Rutherford, R. & Forness, S. (1999) *A Metaanalysis of Social Skill Interventions For Students With Emotional And Behavioral Disorders*. *Journal of Emotional and Behavioural Disorders*, 7, 54-64.
- Stage, S., & Quiroz, D. (1997) *A Meta-Analysis of Interventions to Decrease Disruptive Classroom Behaviour in Public Education Settings*. *School Psychology Review*, 26, 333-368.
- Harden, A. (2003). *Supporting Pupils With Emotional and Behavioural Difficulties (Ebd) in Mainstream Primary Schools: A Systematic Research of Recent Research Evidence of Strategy Effectiveness (1999-2002)* In: *Research Evidence in Education Library*. London: Eppi-Centre, Social Science Research Unit, Institute of Education.
- Scruggs, T.E. & Mastropieri, M.A. (1996) *Teacher Perceptions Of Mainstreaming/Inclusion, 1958-1995: A Research Synthesis*, *Exceptional Children*, 63, 59-74.
- Skidmore, D (2004) *Inclusion, the dynamic of social development*, Buckingham, Open University Press
- Slavin, R. E. (1986) *Best-Evidence Synthesis: An Alternative to Meta-Analytic and Traditional Reviews*, *Educational Researcher*, 15(9), 5–11.