



REVIEW

July 2005

**Strategy training in language
learning – a systematic review
of available research**

Review conducted by the Modern Languages Review Group

The EPPI-Centre is part of the Social Science Research Unit, Institute of Education, University of London

AUTHORS/REVIEW GROUP

Xavière Hassan	The Open University, UK
Ernesto Macaro	Oxford University
Deborah Mason	Oxford University
Gail Nye	Arbor Ridge School, Florida
Pete Smith	The Open University, UK
Robert Vanderplank	Oxford University

Pete Smith
Faculty of Education and Language Studies
The Open University
Walton Hall
Milton Keynes MK6 7AA
p.smith@open.ac.uk
+44 (0)1908 659 353

ADVISORY GROUP

Assistance with reviewing

Concha Furnborough	The Open University, UK
Yasuo Nakatani	Oxford University
Lynn Erler	Oxford University

Searching, document retrieval and administration

Sue Walters,	Oxford University
--------------	-------------------

External Advisor

Lynn Erler	Oxford University
------------	-------------------

CONFLICTS OF INTEREST

Every effort has been made in the review process to be transparent and to follow the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) methodology and quality-assurance procedures for systematic reviewing. As practitioners in the field of language learning and teaching, we acknowledge the possibility that interest and knowledge in the field may have influenced the review in ways which are not apparent to us. Ernesto Macaro and Robert Vanderplank have been involved during the timeframe of this review in primary strategy training research; however, to our knowledge, this has not affected the conduct of the review, and the protocol was published before their research.

ACKNOWLEDGEMENTS

Grateful acknowledgement is expressed to the Open University and Oxford University for institutional support, and to Rebecca Rees, Zoe Garrett, Kelly Dickson, David Gough and Diana Elbourne from the EPPI-Centre.

Funding

Grateful acknowledgement is expressed to the EPPI-Centre and the Department for Education and Skills (DfES) for funding of the review.

The review is funded primarily by finance received from the DfES via the EPPI-Centre, although much of the work on the review is contributed by its authors on HEFCE funded time and, in the case of Gail Nye, her own time during a sabbatical year from teaching.

LIST OF ABBREVIATIONS

DfES	Department for Education and Skills
EFL	English as a foreign language
ELT	English language teaching
EPPI-Centre	Evidence for Policy and Practice Information and Co-ordinating Centre, Social Science Research Unit, Institute of Education, University of London
ESL	English as a second language
ESOL	English for speakers of other languages
HEFCE	Higher Education Funding Council for England
IELTS	International English Language Testing System
L1/L2	First/second language
TLS	Top level structures
RCT	Randomised controlled trial
TOEFL	Test of English as a Foreign Language

This report should be cited as: Hassan X, Macaro E, Mason D, Nye G, Smith P, Vanderplank R (2005) Strategy training in language learning – a systematic review of available research. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

The opinions expressed in this publication are not necessarily those of the EPPI-Centre or the funders. Responsibility for the views expressed remains solely with the authors.

© Copyright

Authors of the systematic reviews on the EPPI-Centre website (<http://eppi.ioe.ac.uk/>) hold the copyright for the text of their reviews. The EPPI-Centre owns the copyright for all material on the website it has developed, including the contents of the databases, manuals, and keywording and data-extraction systems. The centre and authors give permission for users of the site to display and print the contents of the site for their own non-commercial use, providing that the materials are not modified, copyright and other proprietary notices contained in the materials are retained, and the source of the material is cited clearly following the citation details provided. Otherwise users are not permitted to duplicate, reproduce, re-publish, distribute, or store material from this website without express written permission.

TABLE OF CONTENTS

- SUMMARY 1
- Background..... 1
- Aims..... 1
- Rationale..... 1
- Review questions..... 1
- Methods..... 2
- Results..... 3
- Conclusions 6
- 1. BACKGROUND 9
 - 1.1 Aims and rationale for current review 9
 - 1.2 Definitional and conceptual issues 12
 - 1.3 Policy and practice background..... 14
 - 1.4 Research background..... 15
 - 1.5 Authors, funders and other users of the review 15
 - 1.6 Review questions..... 17
- 2. METHODS USED IN THE REVIEW 18
 - 2.1 User-involvement..... 18
 - 2.2 Identifying and describing studies..... 19
 - 2.3 In-depth review 20
- 3. IDENTIFYING AND DESCRIBING STUDIES: RESULTS 25
 - 3.1 Studies included from searching and screening 25
 - 3.2 Characteristics of the included studies 27
 - 3.2 Characteristics of the included studies 28
 - 3.3 Identifying and describing studies: quality-assurance results..... 36
- 4. IN-DEPTH REVIEW: RESULTS 39
 - 4.1 Selecting studies for the in-depth review 39
 - 4.2 Characteristics of studies included in the in-depth review 39
 - 4.3 Further detail of studies included in in-depth review..... 41
 - 4.4 Synthesis of evidence..... 55
- 5. FINDINGS AND IMPLICATIONS..... 62
 - 5.1 Summary of principal findings..... 62
 - 5.2 Strengths and limitations of this systematic review 65
 - 5.3 Implications..... 66
- 6. REFERENCES 70
 - 6.1 Studies included in map (N = 38) and synthesis (N = 25) 70
 - 6.2 Studies excluded from in-depth review, with reasons for exclusion 73
 - 6.3 Studies identified but retrieved after the cut-off (N = 14) 74
 - 6.4 Studies identified after this report’s first draft..... 75
 - 6.5 Other references used in the text of the report 75

Appendix 1.1: Advisory Group membership	80
Appendix 2.1: Inclusion and exclusion criteria.....	81
Appendix 2.2: Search strategy for electronic databases	82
Appendix 2.4: EPPI-Centre keyword sheet, including review-specific keywords .	83
Appendix 3.1: Details of studies included in the systematic map	85
Appendix 4.1: Details of studies included in the in-depth review (N = 25).....	87
Appendix 4.2: Weight of evidence components for studies included in in-depth review (N = 25)	95

SUMMARY

Background

Strategy training in language learning has been topical since the 1970s, and a large amount of work has been done on identifying the strategies used by both successful and less successful learners, and by users of modern languages. The strategies have generally been classified as *metacognitive* (to do with awareness of the learning), *cognitive* (to do with the behaviours and mental process of the learning) and *socioaffective* (to do with personality traits and interactions with others). It is generally held that the skills we develop for learning in general and for learning our first language do not automatically transfer to other learning situations or to other languages – hence the rationale for research into strategies and their potential benefit if the skills are trainable.

Strategy training is defined in this review as any intervention which focuses on the strategies to be regularly adopted and used by language learners to develop their proficiency, to improve particular task performance, or both; a simple example of a strategy in the area of reading skills is brainstorming a theme and making notes in the margin before and during the reading of the text; an intervention that might help language learners to read and understand better.

Aims

The aims of the review are to identify and evaluate the primary research on strategy training in order to gather together, present and comment on the strength of evidence about its effectiveness in teaching and learning of languages.

Rationale

No previous systematic overview of strategy training has been carried out, and literature reviews in the area have not evaluated findings in terms of research type or robustness. At a time when the UK education system, for example, is looking at the possible introduction of strategy training, it is important to get as complete a picture as possible of its effectiveness as revealed through research. This review's systematic searching for study reports, not generally a feature of traditional reviews, aimed to ensure all relevant research evidence was considered, not just that which is most easily accessible. This approach was taken since it is known that the most easily accessible studies may often represent a biased subset of the whole.

Review questions

As its primary question, the review asks the following:

What is the effectiveness of strategy training?

Within this, we also wanted to uncover evidence of differential effectiveness for different languages, different learners (school, university, adult), different stages of learning (beginner, intermediate, advanced), and different language skills (reading, writing, listening, speaking, overall ability, etc). In doing so, we hoped to explore why different types of strategy training might or might not work.

Methods

A peer-refereed protocol, detailing the steps for the review, was published via the EPPI-Centre website. A brief description of key stages is given below, but the main steps in the process include extensive searching for reports of studies, application of inclusion and exclusion criteria to references found, production of a descriptive map of included studies, an in-depth review describing and appraising the quality of a subset of the mapped studies, and a synthesis of findings from these studies. At least two reviewers independently carried out each step of the work, and quality assurance was provided by the EPPI-Centre.

User involvement

The core review group was constituted so as to involve people with experience of research, language teaching at compulsory and post-compulsory levels, and language centre leadership. In addition, consultation events were held with tertiary level students, Government policy-makers and educational researchers expert in the field.

Inclusion criteria

To be included in the map, reports needed to be:

1. of a strategy training intervention in language learning
2. of an intervention carried out in a formal setting such as groups of learners in schools, universities and language centres
3. a study not primarily involving bilingual learners
4. of primary, empirical research
5. of research carried out since 1960

Studies were included in the in-depth review if they met *all* the inclusion criteria for the map and were also experimental in design, testing the effect of an intervention against either another intervention, standard practice, or no intervention.

Search strategy

In summer 2002, 17 electronic databases were searched for studies dating back to 1960 using a range of terms for strategies, strategy learning, or strategy training, in combination with terms for language learning or teaching.

Characterising studies, in-depth review and weight of evidence

To produce the map, studies were screened using the first set of inclusion criteria described above. Relevant studies were classified according to a standardised 'core' keywording system developed by the EPPI-Centre. Review-specific keywording, drawn up by the review group, was also applied so as to describe studies further in terms of the type of strategy training provided and outcomes measured.

Studies included in the in-depth stage of review were subject to examination using EPPI-Centre and review-specific, data-extraction tools. Key elements (such as aims, methodology, context and results) were described and, at the same time, judgments were made about the quality of the reported study. These descriptions and judgments were used to determine a 'weight of evidence' composed of three sections: the trustworthiness of the reported study, the appropriateness of design and analysis for the review question, and the relevance of the focus of the study to answering the review question. A narrative synthesis was drawn up, with each study described and weighted alongside others focusing on similar areas of language learning.

Results

Mapping of all included studies

A total of 567 potentially relevant references were found, and following application of selection criteria, 38 studies (in 40 reports) were finally included in the descriptive map. Most of the studies were on learning English as a foreign or second language (24) but Spanish, French, German, Russian, Japanese, Italian and Latin were also studied. Overall, most of the studies were carried out in the USA (16), but other studies were carried out in Canada, Australia, Hong Kong, Japan, Singapore, Egypt, Germany, Korea, the Netherlands, New Zealand, Poland, Turkey and the UK. Only 11 of the studies involved school students, and the rest were of university, higher education or adult students of languages.

Of the 38 studies in the map, 28 were controlled or randomised controlled trials (RCTs) of interventions, while the rest were descriptive intervention studies, case studies, ethnographic, action research, and one interrupted time series study. Sample sizes varied greatly, and cluster randomisation was frequently used (rather than randomisation of individuals to intervention or control groups). Intervention length also varied greatly from short one- or two-hour awareness-raising interventions to year long programmes of study.

The outcomes measured included accuracy of language output, asking and answering higher order questions, attitude, awareness, comprehension, strategy use, writing ability, and vocabulary recall amongst others. The effects of the interventions on these outcomes were measured using a wide variety of tests and instruments designed specifically by the researchers (interview, survey, self-report, questionnaire, multiple choice and other tests), but in some cases standard end-of-term tests and grade systems were also used.

Studies selected for in-depth review

A subset of 25 of the 38 mapped studies were experimental in design and available in full within the timetable of the in-depth review and these were included in the in-depth review.

Areas of language learning covered by the 25 studies are speaking, reading, writing, overall language ability, vocabulary and listening. Only 10 of these studies explicitly classified the training as focused on metacognitive, cognitive or socio-affective strategies, although the majority (14) were considered by reviewers to be concerned with cognitive strategies or a mixture including cognitive (5). Several of the studies (5) looked at training in semantic and structural mapping of texts (that is, uncovering the meaning and form of a text before or while studying it in detail) in order to improve comprehension. Other interventions involved training in strategies to ask and answer 'higher order' questions to improve overall and speaking ability; mnemonics or keywords for learning vocabulary; focusing on specific grammar items either at the input or output stage; and strategies to improve writing. Of interest to this review is the effectiveness of strategy training, rather than the specific strategies.

Is strategy training effective?

There is sufficient research evidence to support claims that training language learners to use strategies is effective, but it is not possible to say from this evidence whether the effect of training is long-lasting or not. Furthermore it is not really known to what extent the specific mechanics of different training interventions are responsible for the effect, or if it is due to improved awareness that a broad range of training might engender in the learner.

Most studies (N=17) report only positive results. Fewer (N=6) report mixed results, such as reporting a positive finding for one outcome and a negative outcome for another. Only two studies report finding only negative findings.

While most studies report positive findings to a greater or lesser extent, the weight of evidence varies across the different language skills. For speaking ability, for example, training learners to use certain strategies appears successful but the evidence is not compelling (small number of studies, varied relevance, varied reliability) while training in semantic and structural mapping to improve reading comprehension is supported by more robust evidence. Only two studies were found for strategies to improve writing ability (both based on learners revising and/or rewriting first drafts) and both reported positive results – so, despite the small number of studies, the strength of the evidence in them is relatively strong. Interventions to improve overall language ability gave mixed results, and the differences between training interventions plus the design and execution of the studies made it difficult to draw strong conclusions. Similarly, the evidence for the effectiveness of interventions to improve listening ability and vocabulary ability is weakened by methodological characteristics of the studies.

It should be borne in mind that weak evidence does not mean that strategy training does not work – only that the evidence is weak. This is self-evident perhaps, but important as regards the implications: more evidence, or evidence of greater strength, might demonstrate the effect of training more clearly but one should not discount the possibility of findings in the other direction. Only two out of the 25 studies report negative findings, although a further six show mixed results,

with training appearing to work for some aspects of language learning and not for others, or there being no clear difference in outcomes between two strategy training interventions.

Questions that need to be addressed through deeper analysis and by further development of the review include to what extent the effectiveness of strategy training depends on particular aspects of any given teaching/learning situation: for example, the level and stage of learning, language in question, first language of learners, age of learners, prior learning experience, and the similarity of one strategy training context to another.

What sort of studies were found?

All the studies included in the in-depth review are comparative studies looking at one or more interventions compared with another (which might have been an alternative training intervention or a control/usual practice one). Most of them used groups of learners, either in pre-existing classes or randomly formed, but very few of the studies randomised individual learners to the intervention group. Usually the learners were tested before and then immediately after the intervention, and none of the studies retested all the learners again after 6 or 12 months or longer (a few studies did some selective delayed post-testing). Sometimes the effect on the learners was measured through their regular grades or end of term tests, while, in other studies, particular tests were constructed and used. The importance of this is that some of the studies were more 'naturalistic' than others, and this has a bearing on the generalisability of the findings. Some studies related evidence of improvement in performance or ability to increased strategy use, with corroboration of the findings between both performance and frequency of strategy use. There was great variety in the interventions over time and place.

How reliable, relevant and strong were the studies?

The review found that in terms of the total of 27 findings about learning outcomes (two studies reporting on two findings each), three were highly reliable and 15 were of medium reliability – together over half of the included studies' findings – leaving nine of (relative) lesser reliability. This should be seen in the light of all the studies being experimental and a priori more reliable than other forms of research in determining the effect of an intervention on groups of learners (rather than on individuals).

Twenty-five studies were highly relevant to the review question that we were asking. Without prior knowledge of exactly how much research existed, a good number of experimental research studies were found and this provides a convincing example and counter argument to claims that much educational research is not applied and is irrelevant.

The review found that 16 of the studies were considered to be of high or medium weight, and nine of low weight in the evidence they provided: that is to say in terms of overall relevance to the question of the review, appropriateness in their design, and reliability in how they were carried out. This incorporated, in addition to the standardised EPPI-Centre guidelines for evaluation, findings from review-specific questions including indicators in the studies of clear definition of the strategies, clarity in how the strategy is expected to lead to improved learning or

performance, and a relationship between the intervention and the eventual autonomous language learning behaviour of the learners.

Conclusions

Implications for teaching practice

Strategy training for language learning is effective but a number of conditions restrict the universality and usefulness of this finding. Since there have been no direct comparisons, it is not clear whether relatively simple programmes to raise awareness are any more or less effective than longer, expensive, complex interventions. That is to say, it may be possible to make a positive impact on the learner with one or two sessions that ‘open their eyes’ to something. It is unclear what additional impact may be gained through longer interventions that model the behaviours and then seek to have learners imitate and autonomously adopt them. Furthermore, the long-term benefits, or otherwise, of strategy training interventions have yet to be evaluated sufficiently in experimental studies.

Strategy training works for reading comprehension and writing skills, and the research evidence for this is stronger than it is for listening, speaking and overall proficiency.

Practitioners might select from the strategy training interventions found in this review but should also assess carefully their learners and the pedagogic situation in question in estimating the likelihood of applicability and benefit of the outcomes. Additional aspects to consider include whether long or short intervention programmes are required, how similar or different the learners are in relation to those in the experiments, and whether the level and stage of learners are important variables in the particular intervention of interest.

Implications for policy-makers

The evidence shows that it is worth considering strategy training programmes for language learning, on a policy level, as research shows that it is effective in certain situations. As the evidence found is primarily for its effectiveness with adult and higher education learners, more information is needed before straightforward evidence is available for school-level learners. Evidence of a high and consistent standard is still needed, to get a clearer picture in terms for designing policy, and for different levels and stages of learners.

Further evidence should be sought to demonstrate the longer-term effects (say after one year) of training; and research needs to be carried out to investigate the differences and relative effectiveness of awareness-raising and more intensive strategy modelling programmes – these elements have a bearing on the cost effectiveness, both financially and pedagogically, of programmes that might be delivered on a wider scale.

Recent discussions between three EPPI-Centre systematic review groups (Science, Thinking Skills and Modern Foreign Languages) and policy-makers discovered a number of emerging common points that are relevant, including the hypothesis that a combination of behaviour modelling (cognitive strategies) and

awareness-raising (metacognitive strategies) rather than either alone, may be at the heart of improvements observed.

Implications for research

The review has been able to draw some clear conclusions and some equally clear ideas for future research.

Concerning strategy training first, not enough is known about the processes and mechanisms that are in operation, and therefore it remains unclear as to exactly what is making the difference when a strategy training intervention is said to be effective. Future studies should focus on uncovering these mechanisms by tightly controlling the variables in order to isolate the associations between cause and effect, by observing more closely the incidence of strategy use following the intervention, and by seeking in particular to differentiate between the effect of awareness-raising and behaviour-modelling. In addition, the difference between discrete strategies and 'packages' of strategies should also be investigated further.

The research identified in this review has been predominantly with non-school populations, and more evidence of strategy training for school learners of modern languages is needed, particularly if large scale training programmes are to be considered.

Naturalistic evaluations are desirable – that is, studies that attempt to test the intervention with as little disturbance as possible in the day-to-day patterns of the learners in question: this includes using regular tests and grades/scores to measure the effects, and perhaps crossover trials so all participants receive the same treatment overall. If specialised measurement tests and instruments are used, these need to be validated and standardised where possible. It would be good to harmonise approaches across the research community to seek greater synthesis of findings in different studies.

Other enhancements to research designs include carrying out delayed post-testing of the intervention (to see how long the effect lasts) and incorporating case study and ethnographic methods into RCTs to assist with corroboration of overall findings for groups of learners and the effect on individuals within the groups. Clearer randomisation procedures, planned sampling strategies, and better reporting will make the research more reliable and generalisable.

Strengths and limitations of the review

This is the first such systematic review of research evidence in relation to the question. It demonstrates clearly that evidence exists in support of the effectiveness of strategy training in language learning (i.e. that it works), although the caveats constraining this broad statement are important. The review brings to light research evidence that was not previously in the mainstream body of knowledge, and it highlights areas of need in terms of future research, research method and quality.

In terms of limitations, the review looks at an average overall question and does not examine in detail the processes and mechanisms of what is happening in a successful intervention. Meta-analysis of findings has not been attempted and

only 'crude' syntheses of the evidence based on directions of effect were carried out. The evidence needs to be interpreted carefully by practitioners, policy-makers and researchers for specific contexts (e.g. schools versus tertiary sector, beginner versus advanced learner, etc.), and the findings are not immediately transferable to all language learning situations. Furthermore, it was not possible for this review to handsearch journals or to cover many non-English language databases systematically. It is possible that further evidence remains to be identified and included in updates of the review.

1. BACKGROUND

1.1 Aims and rationale for current review

Strategy research in language learning and the training of learners in these strategies have been topical since the 1970s, but the methods for researching this have been diverse, the findings often not incontrovertible, and the need for greater detail or focus clear, although not explicit. At the same time, with a decline in language learning and teaching in many educational sectors in the UK, often excused if not explained by the public at large as a perceived inability to learn languages successfully, the need for discovering or describing effective ways to teach and learn languages is perhaps more sharply relevant than ever.

Within England, the Government has presented its *Framework for teaching modern foreign languages* (DfES, 2003a) and this incorporates clear strategies, if not strategy training, in its objectives (see, for example, Objectives 7W1 on vocabulary, 8W3 on reflective learning, 9W7 on inferencing, 7T1 on structural mapping in reading comprehension, 8L4 on communication strategies, and 7C5 on social and linguistic conventions) yet no comprehensive or systematic overview of the evidence base for this has been carried out to underpin its introduction. There is a vast expertise and a critical mass of tacit knowledge in the field, but as yet there has been no structured overview of it based on research typologies and associated reliability.

The Government's 2004 consultation document for Key Stage 2 and the teaching of languages at primary level in schools explicitly covers language learning strategies as an objective:

- 1.14 **Language Learning Strategies** are an important part of the suggested Key Stage 2 programme. By having regular opportunities to identify and apply such strategies, children will become more aware of how they learn languages and should become better language learners. The Language Learning Strategies are also integrated with the objectives for Oracy, Literacy and Intercultural Understanding (DfES, 2004, section 1.14).

The need for the review stems primarily from the absence of any systematic overview of the research into strategy training for language learners. Literature reviews (for example, McDonough, 1999) do not describe the criteria by which the research is selected, and do not take into account the reliability of the findings as a function of the research method (although, in fairness, the example cited did not claim to be a systematic review of effectiveness).

Furthermore, the literature reviews which exist are all paper publications and do not include updates or contemporaneous feedback and reaction; this is significant given the amount and quality of recent and ongoing research.

The aims of the review are as follows:

- To review systematically the research evidence on the effectiveness of strategy training in improving proficiency in language learning

- To descriptively map identified research into strategy training
- To assess the quality of research into strategy training

This review will be updated on an annual basis with reports of primary research incorporated as and when they are identified or published.

1.1.1 Strategy training context

This review defines a learning strategy as any intervention which focuses on the strategies regularly to be adopted and deployed by learners in order to develop language proficiency, improve language task achievement or both. This incorporates the essence of the definitions found in the research literature: for example, Chamot (1987), Cohen (1998) as reported in Macaro (2001), Oxford (1990), Weinstein and Mayer (1986), and Wenden and Rubin (1987).

The sense in which it is used in this review is to cover activities or programmes that explicitly set out to equip language learners with 'learning tools' prior to embarking on (or during) their principal course or programme of language study. In its simplest form, it might be two or three sessions, prior to starting the course work proper, on how to learn vocabulary, or what it means to learn grammar, or discussion of definitions of communicative performance versus competence. A more complex model might be a package of strategies, either before or during a course, that the learner develops and maintains for him or herself, assessing progress as the course goes on (self-evaluation and reflection), and adjusts learning style and strategies to the tasks in hand.

All language learning (if not all learning) involves a degree of change in awareness as well as language proficiency: it seems virtually impossible to take a language course and not learn something about the learning process on the way, and it is likely that even the least successful learners do so (O'Malley, 1987). However, this review focuses on research attempts to achieve this proactively, although it is necessary to take into account both the explicitness of this awareness, and differences in awareness that may already exist at different stages in a student's progress.

Advances in our understanding of cognition in general and the relationship between working memory and long-term memory in particular (Miyake and Shah, 1999; Eysenck and Keane, 2001) allow us to explore with much greater direction and purpose the processes involved in acquiring a foreign language. We are becoming increasingly aware that underlying these processes are a series and range of learning strategies and a considerable body of descriptive research bears testimony to this, although prospective, empirical, experimental studies have not necessarily matched it.

One of the foremost promoters of strategy training, Barbara Sinclair, at a British Council conference in Oxford in July 1999, started a presentation by stating that even after ten years of strategy training, she did not know if it worked (Sinclair, 1999). As far back as 1969, researchers expressed doubts about the effectiveness of strategy training and claimed amongst other things, that the learners rarely see the relevance of what they are doing (Benson, 1995; Carroll, 1973; Carton, 1971; Politzer and Weiss, 1969; Rees-Miller, 1993; Smith, 1985). Much of the rationale for practising strategy training is theoretical and ranges from the instinctively attractive proposition that good learners are independent to the

plethora of ways in which practitioners claim to go about achieving it. Dickinson (1992) states, 'Few teachers would seek to make learners dependent on teaching for evermore; most of us recognise that the ability to learn independently is a proper outcome of teaching...' (p 2), although Macaro (1997) is more circumspect, stating '...there is no evidence as yet of a link between increased strategy use and increased language competence' (p 119). However Macaro (2001) finds sufficient evidence for more optimism that a link exists and his ongoing research in the field (Macaro *et al.*, 2003) will contribute further evidence one way or another.

Strategies, mostly in the area of English language teaching (ELT), have been theoretically or non-empirically described and designed for a variety of learning needs, including reading (El-Koumy, 1997; Kern, 1989; Singhal, 1998; Zhicheng, 1992); writing (Gooden-Jones, 1996; Kresovich, 1990); speaking (Ishii and Klopff, 1980; Luk, 1992); listening (El-Koumy, 1997; Viswat and Jackson, 1994); global language ability (Chamot and O'Malley, 1996) or discrete areas such as grammar (Jo, 1997) and vocabulary (Kaelin, 1991; Weatherford, 1990).

The strategies may be implemented in a number of pedagogical formats or modes: for example, co-operative learning (Correa, 1995; Gooden-Jones, 1996), awareness-raising (Yang, 1995) or via academic study skills and access and foundation courses.

A variety of means of assessing or measuring the strategies was used, including collaborative feedback (Lane and Potter, 1998), video-filmed assessment of performance, learner profile questionnaire (Oxford, 1990), negotiated syllabuses and process-oriented tasks.

A recurrent strand to strategy training is concerned with learners gaining greater language awareness, or metacognitive awareness, and a subsequent self-initiated deployment or use of the strategies leading to greater autonomy and independence (Baird and White, 1984; Moulden, 1981; Yang, 1995).

Some of the strategies are self-learned or self-taught while others, or alternative modes of delivery involve taught elements – that is, strategy training can be implemented as a self-directed or a taught component of a programme. Both are of interest in this review if the underlying intention is proactively to introduce or nurture an adoption of learning strategies by the learners.

However, much of the justification for strategy training is opinion, selective research, personal experience, theory and speculation (Biggs, 1987). While the rationale for learners having more autonomy in their learning is intuitively desirable, and much of the theory rationally and persuasively underpinned, there is no comprehensive reliable overview of research evidence that shows either that successful learner outcomes are related to such strategies, or that unsuccessful learners could learn them or could use them, and that they would have the desired outcome (Harris, 1997). There is a need for a survey of the quantity and breadth, and an evaluation of the quality of such research as is available, and an attempt to bring it together to obtain a more comprehensive, transparent, reliable and applied knowledge base.

So while the use of strategy training is based on theory as much as sound evidence, research continues (Chamot and O'Malley, 1996; Oxford, 1986; Rubin, 1975, 1981; Starks-Martin, 1996) into the strategies used by both successful and less successful language learners, on the assumption that, by identifying and analysing such strategies, and then developing them into teachable or learnable

pieces, learners can benefit. However, even if this is the case, the mechanisms are not known with much certainty (Perez, 1990). Additionally, the success of training may well be entirely dependent on the learning and cognitive style preferences of the individual (Wenden and Rubin, 1987).

Language learning and teaching in schools in the UK at the beginning of the millennium are in recession in a number of ways. Teachers are leaving the profession, recruitment of language teachers is difficult, and the number of students taking modern and foreign languages at exam level at school (and therefore at university) is dropping significantly. This situation is all the more disturbing as the UK is perceived to be standing in the wings of Europe while its continental neighbours forge ahead in economic and monetary union where language skills are key tools (Government Green Paper, 2002). The problem is recognised, and for England the DfES (Department for Education and Skills) National Languages Strategy is forthright in its intention to respond:

Changing the country's attitude to teaching and learning languages will demand a huge cultural change. It will rely on action from Government, schools, LEAs, colleges, universities, employers, parents and learners. Partnership is the key to making this strategy a reality. It is only if we work together on the implementation of this strategy that we will achieve real and lasting change for the future, for our young people, for adults, for business and for our society. (DfES, 2003b, p 5)

The Nuffield Enquiry report (Nuffield, 2000) has made plain in its recommendations many of the areas that might be given attention, and recent Government initiatives aim both to improve the recruitment and training of teachers, and to increase the number of students taking foreign and modern languages.

Much work has been carried out on learning strategies, but previous reviews of strategy training are either not systematic or comprehensive (McDonough, 1999) or are not primarily reviews of effectiveness (O'Malley and Chamot, 1990; Weinstein and Mayer, 1986). Lack of valid and reliable evidence is a problem that besets our knowledge of the effectiveness of strategy training, and a considerable amount of what poses for theory-driven research within education is in fact merely opinion, and this is often reflected in traditional literature reviews. A systematic review of research evidence (and not just a selection of some evidence), carried out and maintained on the basis of a transparent protocol that specifies outcomes, interventions and research methodology as vital elements, may go some way to providing a more reliable synthesis on which practitioners and individuals can base decisions. Within the context of such a systematic review, and particularly concerning the modes of delivery and processes involved in strategy training, descriptive research will play its part.

1.2 Definitional and conceptual issues

1.2.1 Definitions of strategies

Since the start of research into strategies in the 1970s, an accepted framework for describing them has emerged based on the work of Rubin (1975) and Stern (1975) followed by work from O'Malley and Chamot (1990) that have crystallised

into taxonomies of strategies (see, for example, Ellis and Sinclair, 1989) that have served until now in providing the basis for description in the following tripartite way:

1. Metacognitive strategies, such as advance preparation, analysing needs, comparing, expressing beliefs, prioritising, setting short-term aims, monitoring, evaluating
2. Cognitive strategies including defining, inferencing, keeping a diary, listening for gist, predicting, reading aloud, skimming, translation
3. Affective/social strategies, such as discussing, joining a group, motivation, attitude

This perspective on strategies, based on much empirical (although not experimental) research, has provided the theoretical framework within which most strategies training research and development have taken place.

Despite a general acceptance of this traditional framework, it was perceived that a tighter definition of strategies was needed for this review. The autonomous use of strategies is not necessarily included in a traditional definition of strategy training but was considered an essential element of this review. Strategy training interventions were therefore defined for the purposes of the review as interventions that set out to train learners to notice, and then do something in order to improve an aspect of their ability to learn the language. This was built into the guidelines for assessing the studies with a view to differentiating more precisely between strategies and what might only be teaching or learning methods.

To distinguish between teaching methods and strategy training, the following guidelines were used:

- Have the strategies been defined (if X and Y, then Z) by the researcher/teacher?
- Is the way the strategy is supposed to lead to learning or improved proficiency clear?
- Is there a clear relationship between the strategy training (what the teacher did with the students) and what the students would be expected to do eventually as independent individual learners?

As described in section 2.3.1, this definition of a strategy training intervention was developed after the initial screening of studies for inclusion in the review's descriptive map, at the stage of screening studies for inclusion in the in-depth review. (Studies had initially been recognised as being about strategy training by reference to a strategy in the title or abstract of the report.) It was then applied retrospectively to all studies initially included in the map to check that the distinction between teaching methods and strategy training had been made.

The review is less concerned with such learning strategies, or the relative merits of strategies, than with the effectiveness of training students to use, and then deploy autonomously, the strategies – whatever they are. Such interventions might be training students how to plan and organise their study consciously, how to improve their reading comprehension, how to guess the meaning of unknown

words, or training them to write better essays – but crucially the interventions must include the element of training and not merely be teaching of the outcome.

1.2.2 Additional issues in defining strategy training

Education and psychology rather than specifically language?

A query of a more general nature concerns the distinction that may be drawn between strategies, skills and behaviours that are unique to modern foreign language learning and those which are non-language specific. The difference is not obvious, and a pragmatic acceptance has been adopted for this review in as much as the setting (that is, language teaching and learning) is the factor which makes the intervention relevant, even if the skills can be deployed by learners in other areas of their education and lives.

Problem of assessing training effectiveness but necessarily testing the strategy at the same time:

This review intended to focus less on the particular strategies than on the effectiveness of training of learners overall in the strategies whatever they are (future reviews could usefully examine the effectiveness and mechanisms of particular strategies). However, this raises the possibility that theoretically we might be dealing with the effectiveness of training – but in strategies that don't work, which a priori is not a very relevant or useful endeavour. Furthermore, it would require a complex study design to test such hypotheses and none set out to do this: that is, the studies only look at the outcomes of the training as functions of the strategies, not as functions of the training. We accept that the effectiveness of the training is tied in closely with the effectiveness of the strategies themselves, but that association of the use of the strategy with learning outcomes gives at best proxy evidence of the effect of the training.

1.3 Policy and practice background

1.3.1 Evidence-based policy

It would be reasonable to assume that, with so much research on strategies and their importance in language learning and teaching, there should already be a clear picture of the state of affairs regarding its mainstream use in schools, universities and language learning in general.

However, this is not the case and, despite the research, the perceived value of strategy training is still very much based on opinion, received wisdom and claims of good and best practice. A kind of orthodoxy has grown up around strategy training, and apart from a period of time when learner training became fashionable and conjoined with process-oriented syllabuses in ELT, there has been no systematic incorporation of strategy training in the mainstream.

In this context, allied with the maturity of the field, and with an increasing will to base practice on reliable evidence, an overview of such research into its effectiveness is desirable and should be available.

The current decision-making climate calls for increased reference to research and its use as a basis for policy, and, despite reservations in certain quarters about both the theory and practice of this, it looks set to continue for the time being. The language teaching profession is to some extent becoming more research-based, and can become more evidence-informed in a number of ways, while still maintaining independence of thought. According to Pachler (2003) for example, researchers can

provide research training to practitioners interested in engaging in evidence-based practice; focus their research on areas perceived to be of relevance by practitioners; write the findings up in a way that engages with the discourse conventions used by practitioners and disseminate them in forums practitioners readily access; and play a vital role in synthesising and summarising existing research evidence to make it more readily accessible. (Pachler, 2003, p 10)

This review addresses these needs.

This is perhaps particularly necessary in the UK school domain following the Nuffield Languages Report (2000), current concern at the falling numbers taking or teaching languages, and recent changes by the Government to policy on languages in schools.

1.4 Research background

Wider picture

No systematic reviews of strategy training have been identified, but two overviews or literature reviews exist. There is a considerable amount of quasi-experimental and descriptive research, and an abundance of opinion reports and articles which are regularly cited as research.

No standard approaches

The outcomes of training (competence, performance, learning, etc.) in the research studies are not measured by standardised tools or means and sometimes proficiency is an outcome of interest, sometimes not. The effect of interventions is not assessed by delayed and long-term post-intervention testing. A significant amount of research has been carried out in the area, but without any harmonisation in approach, and aggregation of findings across the field is not straightforward. Some of the research identified for this review has apparently not been not cited elsewhere, yet contains extensive research data and evidence.

1.5 Authors, funders and other users of the review

1.5.1 Users of the review

The initial motivation for this review was the contact author's desire as a practicing teacher of languages to find out whether it was justified or not systematically to incorporate strategy training in language teaching and learning.

Discussion, and the research identified on an ad hoc basis did not provide a straightforward answer.

The review is intended for a number of different end-users, but ultimately it is for learners of foreign languages who stand to benefit most directly from any reliable knowledge originating from research. If it is clear from research (and other evidence, as a review cannot definitively answer its question and there will always be caveats) that strategy training is effective, then the profession can inform its policy and practice decisions with greater confidence. Other questions will undoubtedly emerge as the picture becomes clearer, surrounding such issues as the kind of strategy training for specific circumstances and for whom. Notwithstanding these, teachers of languages and researchers of means and methods will be able to make use of the review.

In addition to learners and teachers, wider benefits will accrue to parents, curriculum and materials designers and policy-makers, and versions of this review for different areas of interest are expected to be made available.

1.5.2 Authors

The authors of the review all have an interest in the knowledge that the review can uncover and present, and as such they represent the various parties likely to use the review.

Ernesto Macaro, a published author and researcher in the area of learner strategies, is a teacher trainer and director of the Applied Linguistics master's degree at Oxford and is currently conducting experimental research into strategy training. Robert Vanderplank, Director of the Language Centre at Oxford University, is concerned amongst other things with undergraduates' ongoing language skills maintenance; Deborah Mason, Assistant Director of the Language Centre at Oxford University, teaches English for academic studies to postgraduates. Both are concerned with policy and management. Peter Smith and Xavière Hassan are lecturers in French at the Open University and directly involved in devising materials for, and teaching adult learners of, languages. Gail Nye is a teacher of English for speakers of other languages (ESOL), and learner and user of Spanish in Florida, USA.

1.5.3 Policy

The review is timely, not only in relation to the Nuffield Languages Inquiry Report (2000) and its recommendations (14. 'Ensure policy is reliably and consistently informed', and particularly 14.3 'specify key areas for attention in language teaching and learning, such as autonomous learning...', p 97) but also in relation to research into strategies (see, for example, the strategies referred to in the Key Skills Stage 2 descriptions earlier) aimed at making language learning more effective. While it is generally accepted that successful learners use strategies, the perceptions of the strategies by learners themselves are not necessarily comprehensive or well-developed (Cajkler and Thornton, 1999).

The Government's 14–19 Green Paper (the section on languages) also highlights the need to raise standards in language teaching and learning, although according to some it does not spell out specifically how it might be done (Pachler, 2002). It does aim, however, to establish specialist language colleges, and this

review is relevant to such achievement of higher standards – not least as it looks at strategy training across any and all modern languages rather than particular or a limited range of languages. With specialist language presumably expected or intended to be focal points of good or best practice, it is planned to involve a school in consultations on the findings of this review.

1.6 Review questions

The primary question addressed by the review is:

What is the effectiveness of strategy training?

It also addresses the following set of sub-questions:

1. *Does strategy training work*
 - *for all language skills (reading, writing, listening, speaking, overall proficiency, grammar ability, motivation)?*
 - *for all learners?*
 - *for all languages?*
 - *at all stages of language learning instruction?*
2. *If it appears not to work, what might be the reasons?*

As the first stage of this review, a descriptive map of relevant research was produced. The aim was to provide a broad overview, both comprehensively and systematically, to identify the research available and point to areas where there are gaps in the knowledge/evidence base.

The review's protocol also outlines that a second, in-depth stage of the review would focus only on studies using some sort of comparison group design, if sufficient studies of this type were identified.

2. METHODS USED IN THE REVIEW

2.1 User-involvement

2.1.1 Approach and rationale

The approach to user-involvement for this review was to incorporate, through the authors becoming the reviewers, different aspects of experience and interest in language learning and teaching. This has applied from the planning and preparation of the review, throughout the different stages, and in the drafting of the findings.

The authors – five out of six of whom had no prior experience of systematic reviewing, and only one of whom had specialist knowledge of strategies and strategy training – brought in relevant experience, as potential users of the review, in the following ways:

- language learning (all six authors are learners and users of modern languages)
- parents of children learning languages at school (four of the six authors)
- language teaching (all six authors are or have been language teachers, covering young learners at school and tertiary level, and adult learners)
- policy-making (one director of a language centre, one assistant director)
- researchers (four of the six authors)

2.1.2 Methods used

All the authors – all modern language learners, users and teachers – were involved in all stages of the review.

In addition, at the in-depth evaluation of included studies, a study evaluation day was organised with a group of students studying for their master of arts (MA) degrees in applied linguistics. This was aimed at introducing additional users to the process and also expediting the study evaluations. Two of these students (one a practising lecturer in Spanish and researcher, and one an educational researcher) completed data extractions of a study, working as the second reviewer with one of the report authors in both cases.

2.2 Identifying and describing studies

2.2.1 Defining relevant studies: inclusion and exclusion criteria

To be included in the map, reports needed to be of

1. a strategy training intervention in language learning
2. an intervention carried out in a formal setting such as groups of learners in schools, universities and language centres
3. a study not primarily involving bilingual learners
4. primary, empirical research
5. research carried out since 1960

2.2.2 Identification of potential studies: search strategy

Reports were sought via database searches and by contacting colleagues in the field.

Key databases were identified and a draft search strategy was developed during January and February 2002 while searches up to the cut-off date were run from June to September 2002. The databases searched and some of the terms used to build database searches are listed in Appendix 2.2. It was decided to search for reports that referred to

- (i) strategies, strategy learning, or strategy training
- (ii) language learning or teaching

Searches used database controlled terms, free text searches, or both, depending on the individual database and the availability of database controlled terms and thesauruses. The searches were recorded, and a quality check of the searching process conducted afterwards (see section 2.2.5).

No systematic effort was made to identify relevant studies in the non-English language research literature, although any non-English language reports found were included in the review process.

2.2.3 Screening studies: applying inclusion and exclusion criteria

The inclusion/exclusion criteria were first applied independently to each of the abstracts and/or titles of reports by two reviewers. A list was drawn up indicating whether to include or exclude according to the criteria set out in section 2.2.1. This list was scrutinised further against the inclusion/exclusion criteria at two separate meetings of the review authors. After this, full reports were sought for studies marked for inclusion. The cut-off date for retrieval was set as 21 May 2003.

Once obtained, full reports of studies were screened again, using the same criteria. At this point, additional guidance on defining strategy training was developed, as described in section 2.2.1. This final screening for the map was conducted by one reviewer (PS) who consulted at least one of the review authors when proposing that a study be excluded. Excluded studies were retained for background and supporting material.

2.2.4 Characterising included studies

Full reports were obtained and classified according to a standardised 'core' keywording system developed by the EPPI-Centre (EPPI-Centre, 2002a). This classifies studies in terms of the type of study; the country in which the study was carried out; the educational focus of the study; and the study population. Reports also were classified by the review group so as to describe further the study type, size, language skills addressed, strategy training type, language learned and outcomes measured.

2.2.5 Identifying and describing studies: quality-assurance process

Some known references did not appear in the database search yields. Therefore checks were carried out across the searches to gauge their effectiveness. A total of 26 reports by known authors that had not been identified by the searches were tracked to determine why their studies had not been identified.

The application of inclusion/exclusion criteria and the allocation of core keywords were both independently considered by a member of the EPPI-Centre (RR) for a total of 19 reports and queries fed back to the review group. The lead author also checked the application of core keywords for all reports finally included in the map.

2.3 In-depth review

2.3.1 Moving from broad characterisation (mapping) to in-depth review

In establishing the criteria for which studies to include in the in-depth review, the team considered their review question and balanced the need to focus on research, such as large scale studies that control for various sources of bias, and other descriptive research that also forms part of the overall picture for the purposes of policy considerations.

An additional criterion was drawn up to identify studies from the map that would be reviewed in-depth (see criterion 6 below). In short, to be included in the in-depth review, studies needed to meet *all* the following criteria:

1. a strategy training intervention in language learning
2. an intervention carried out in a formal setting such as groups of learners in schools, universities and language centres

3. a study not primarily involving bilingual learners
4. primary empirical research
5. research carried out since 1960
6. experimental studies testing the effect of the intervention against another intervention, or standard practice, or no intervention

Screening for the in-depth review was performed by two reviewers. These reviewers worked independently of each other and then met to reach consensus over inclusion and exclusion where necessary.

The protocol for this review originally proposed that studies would only be included in the in-depth review if the strategy training intervention under evaluation was a minimum of two hours in duration. It was felt that nothing less could have a genuine effect on overall language learning proficiency. At the stage of screening studies for inclusion in the in-depth review, through discussion with members of the EPPI-Centre and within the group, the decision was taken to include studies of interventions shorter than two hours. On balance, it was agreed that moments of epiphany, or the discovery of something one did not previously know, can be of considerable 'life changing' value for a learner and the impact of an intervention, therefore, is not necessarily a function of its duration. The initial criterion was determined to have been untenable, and interventions to raise awareness might have been unjustifiably excluded. This modification to the protocol resulted in the inclusion of two further studies.

At this stage, the group also identified aspects of studies that would be included that were of particular interest to explore further through inclusion in the in-depth review and subsequent data extraction; these relate to types of participant, types of educational setting, and outcomes.

Types of participant

Learners at different stages of development (young learner, school, university, adult) were included in the selected studies (as long as they were not bilingual or on their third or subsequent language as far as could be determined from reports). For the purposes of the in-depth review, learners were grouped according to different parameters, including age, language and level of proficiency. Learners of languages who are living or resident in the country or culture whose language they are learning (that is, learning Spanish in Spain rather than in Canada) were included. It was considered likely that a lot of the available research might be in the area of English language teaching (ELT) or English for speakers of other languages (ESOL) and so it was anticipated that sub-grouping of the populations could be carried out as a guard against reaching potentially confounded findings.

Types of educational setting

A major motivation as described earlier for the review was to identify the evidence relevant to the UK school situation, and, where possible, findings would be grouped to inform this and other specific educational settings.

Outcomes

The review was interested specifically in studies that looked for improved proficiency on one or a number of measures as a result of the strategy training

provided for the learners. This could mean overall learning proficiency as measured in regular examinations and assessments, or specific areas of improvement such as vocabulary acquisition, accuracy in oral expression, grammar ability, reading comprehension or a number of other measures – whether or not they are also proxies for overall ability. It was felt that self-report indicators of change or improvement were also of interest, but measures of triangulation would be desirable in such cases to counter the well-known problems associated with self-report. In addition, independent measures of strategy deployment, regardless of proficiency or ability indicators, were of interest and deemed essential to identify measurement of change after a period had elapsed post-intervention. Given that the notions of effective strategy training and the nurturing of independent (autonomous) learning are mutually dependent in language learning, it seemed self-evident that the effect of a successful intervention should not disappear rapidly after the end of the experiment, so delayed post-testing was incorporated as an outcome of interest.

2.3.2 Detailed description of studies in the in-depth review

Data were double-entered into EPPI-Reviewer (software) by two reviewers working independently on each included study. In cases where there was initial disagreement about data extraction or quality appraisal, this was discussed and resolved. Two members of the EPPI-Centre (RR and ZG) were involved in this process.

A standardised data-extraction framework was used: the EPPI-Centre's review guidelines for extracting data and quality assessing primary studies in educational research (EPPI-Centre, 2002b). For each study, data were collected on the aims, study sample, recruitment, data collection and analysis methods, findings and authors' conclusions. An additional set of guidelines, drawn up by the review group was used to classify studies further in terms of the type of strategy training provided and outcomes measured (see Appendix 2.2).

2.3.3 Assessing quality of studies and weight of evidence (WoE) for the review question

The quality of the studies was evaluated using questions contained in the EPPI-Centre's review guidelines, referenced above, with the aim of assessing to what degree readers could depend on the reported study findings to address the question of the review. This took into consideration the following:

- A: Soundness of method (the extent to which a study is carried out within the terms of that method – this concept is referred to in the rest of the report as reliability)
- B: Appropriateness of study type to answer the review question (appropriateness of methods to the review question)
- C: Relevance of the study focus to the review question (addressed by examining questions about the study's stated aims and rationale, research questions and its policy or practice focus, and comparing these with the areas of specific interest to this review using the findings from the additional validity and reliability review specific questions – see section 2.3.1 and Appendix 2.2).

The three criteria above were taken together to provide an overall compound evaluation of the weight of evidence attributed to the results of each study in relation to addressing the review question. Each of the criteria was classified as high, medium or low, and these were considered together in the light of the individual characteristics of each study, and particularly with regard to generalisability of the findings (one aspect of WoE A), to provide an overall compounded evaluation (WoE D). Overall studies could not obtain higher for WoE D than they had done for WoE A, but in making this judgement relevance (WoE C) was also taken into account.

2.3.4 Synthesis of evidence

Frequency and cross-tabulated reports were run on the studies. Study findings were grouped according to the following two broad bands of criteria: descriptive characteristics and evaluated findings.

Descriptive characteristics

- Study type
- Sample size
- Year of study
- Publication type
- Language skills studied
- Intervention type
- Intervention length
- Outcomes measured
- Education sector
- Country of study

Evaluated findings

- Speaking ability
- Reading ability
- Writing ability
- Listening ability
- Overall language ability

Findings might have been grouped in a number of different ways, the 'bluntest' approach being to put them all together (or 'lumping') and seeking to comment on all the studies as exponents of strategy training in the broadest sense. However, this was considered undesirable due to the many differences between studies along the parameters referred to above.

Consequently, the justified degree of meaningful 'lumping' was considered to be in terms of domain or outcome: that is, speaking, reading, writing, listening and overall ability where there is a greater uniformity of metric (that is, measurement of, say, reading comprehension). This does of course limit the number of studies that might be combined in pursuit of an aggregated assessment of evidence, and questions of whether this compares like with like may still arise for some.

At this stage of the review, numerical data were not extracted and meta-analyses were not conducted. An update of the review intends to consider the possibility and desirability of such meta-analysis.

2.3.5 In-depth review: quality-assurance process

An EPPI-Centre staff member (RR) was present at an initial meeting where the review group completed data extraction as a team on one study. All data extractions were conducted independently by two reviewers who then met to reach consensus. For 11 of the final data extractions, one of the joint reviewers was an EPPI-Centre staff member (RR and ZG).

3. IDENTIFYING AND DESCRIBING STUDIES: RESULTS

This chapter reports the results of searching for and screening studies for this review. It then describes in detail the characteristics of studies found and included in this review's systematic map: empirical studies of strategy training interventions in modern language learning (not primarily involving bilingual learners) in formal settings and conducted since 1960.

3.1 Studies included from searching and screening

A total of 567 potentially relevant references were found through the searches described in section 2.2.2. Figure 3.1 summarises the number of studies at each stage of the review. Full hard copies were sought through library loans or other means for 83 reports. Of these, full text reports for 74 studies were obtained before the retrieval cut-off date. Of the 74 reports retrieved, 20 failed to meet the map's inclusion criteria and so were excluded. Fourteen reports were retrieved but after the review's cut-off point so were not considered further; these are listed in section 6.3. A total of 38 studies (contained in 40 reports) were included in the map.

3.1.1 Search and screening results as an indication of the broader research base

A large amount of research on language strategies was found during searching and screening. From the research studies that we excluded (data not presented), we estimate that the research included in our map constitutes less than 10% of the total number of research reports on language strategies found. Studies that were not included in our map, but that may be of value include:

1. studies of strategies or strategy use that do not report on strategy training per se
2. studies in informal instruction settings
3. research that is theoretical rather than empirical

3.1.2 Identifying and describing studies: quality-assurance results

Of the 26 potentially relevant reports checked for an explanation of why they were not found in database searches, 13 did not appear in any of the databases that had been searched. Of the 13 that did appear in the databases when searched for under author name, the reason for not identifying them were as follows:

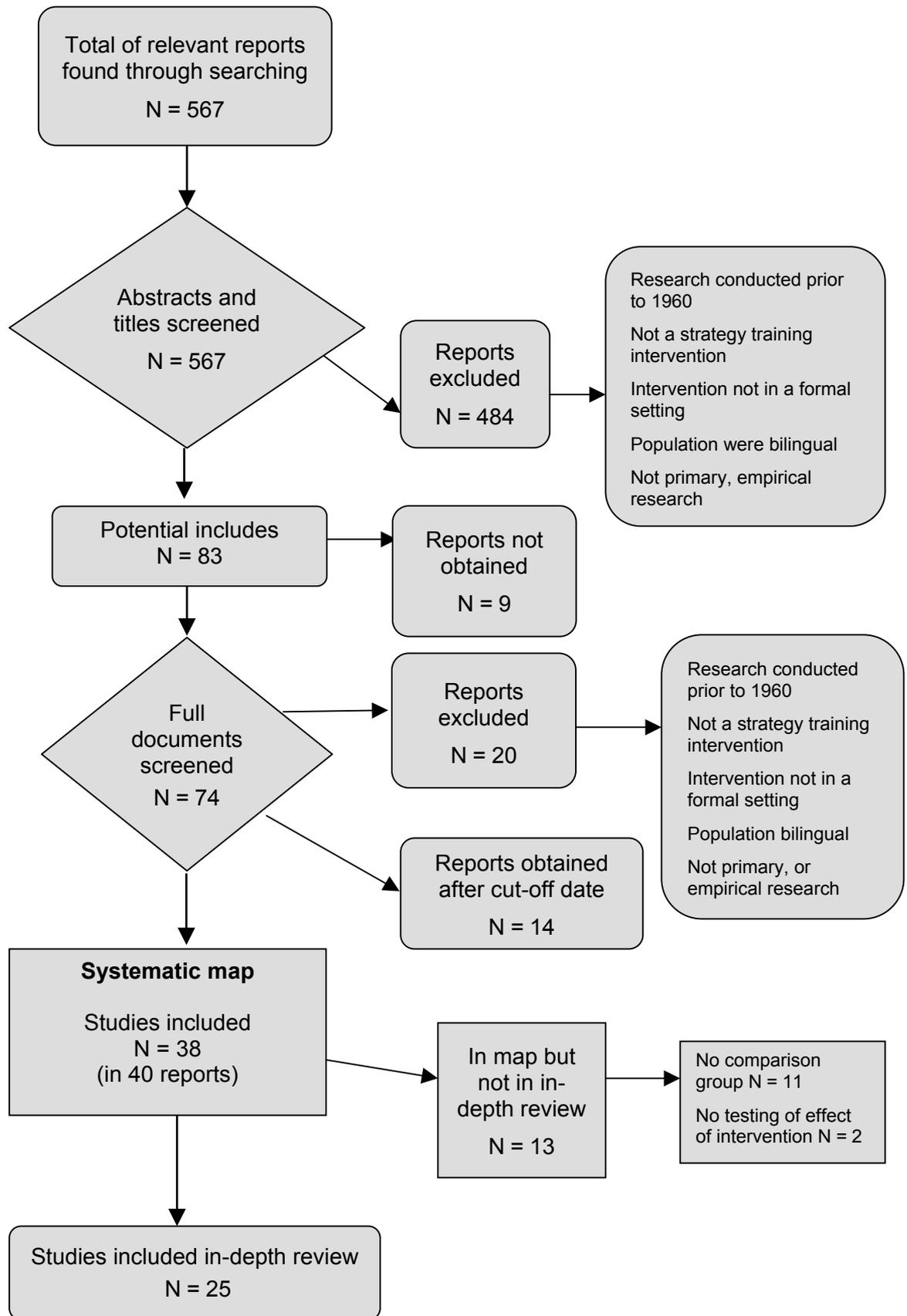
- Idiosyncrasies of searching one or multiple databases at the same time (N=4): for example, different results on different databases when same search term or operator is used

- Indexing of the report does not identify it as modern or foreign language research (N=4)
- Reviewer error in excluding wrongly (N=4)
- Possibility that the reference was added to the database post-search (N=1)

The studies found through these checks (N=13) were added to the total that was screened for inclusion.

The independent screening conducted by the EPPI-Centre identified difficulties in applying inclusion criteria and this led to the production of a tighter definition of strategy training (as described in section 1.2.1). This revised definition was then reapplied to all studies by the lead author for all initially included reports.

Figure 3.1 Filtering of papers from searching to map and in-depth review



3.2 Characteristics of the included studies

The map includes a total of 38 studies conducted since 1980 and reported in a total of 40 reports. The two studies reported in more than one report are cited in this review as Najjar (1997), and Lawson and Hogben (1998) (see section 6.1 for details). The following sections describe a number of key aspects of the 38 studies.

3.2.1 Study dates

From the start of 1996 to the end of 2003, 26 studies were reported. Prior to that, going back to 1980, only 12 studies were completed. The mid-90s therefore appear to mark the beginning of an increase in the number of studies: on average, 0.8 compared with 3.25 studies per year for the two periods respectively (that is, greater than a fourfold increase). Dates of the studies have been recorded in Table 3.1 where known, or date of publication minus one year where the date is not explicit in the report.

Table 3.1 Studies by year (N = 38 studies)

Date	N	Running total
1981	1	1
1985	1	2
1989	1	3
1991	1	4
1992	2	6
1993	3	9
1994	1	10
1995	2	12
1996	5	17
1997	5	22
1998	2	24
1999	6	30
2000	5	35
2001	2	37
2002	1	38

3.2.2 Study size

Over all the studies, a total of over 3,000 learners of languages participated, with sample sizes ranging from 1 to 863 (see Table 3.2). The majority of studies have sample sizes over 30, but this is not necessarily an indicator of robustness considering that many of them were cluster randomisations (whole groups, perhaps only two or three classes – data not shown) and perhaps only large samples of over 200 are more realistic for cluster randomisation trials in terms of assuring the power of the sample to detect an effect reliably. In some studies, individuals rather than groups were randomised.

Table 3.2 Sample size in studies (N = 38 studies)

Band	Sample sizes (participants in studies)	Number of studies
1–5	1, 4, 5	3
6–10	7, 8, 10	3
11–20	20	1
21–30	21, 26, 28	3
31–40	30, 32, 32, 32, 34, 36, 40,	7
41–50	43, 45, 48	3
51–60	51, 58,	2
61–100	68, 71, 91, 75	4
101–200	108, 119, 141, 143, 151, 158, 187	7
>200	229, 244, 338, 863	4
Unclear		1

3.2.3 Study type

As Table 3.3 indicates, the majority of the strategy training studies in the descriptive map used experimental comparative/controlled study designs: 28 out of the 38 were comparative or randomised comparative studies. Some had control groups receiving no intervention or participants receiving their regular language courses, while others were comparisons between two or more alternative interventions.

Table 3.3 Study types used to test strategy training interventions (N = 38 studies)

Type	N
RCT (randomised controlled or comparative trial)	15
Comparative study (controlled trials that did not use random allocation)	13
Descriptive intervention study	6
Case study	1
Ethnography	1
Action research	1
Interrupted times series	1
Total	38

Note: Study design categories are mutually exclusive.

3.2.4 Language skills

As Table 3.4 indicates, the majority of the studies examined reading, but nine principal aspects of language skill were covered altogether.

Table 3.4 Language skill studied (N = 38 studies)

Skill/domain*	N
Reading	12
Vocabulary	6
Listening	6
Speaking	4
Overall language ability	5
Strategy use	4
Writing	3
Awareness	3
Grammar	1

Note: A single study could cover two or three skills.

3.2.5 Intervention type

Table 3.5 presents the training interventions described in terms of whether they targeted cognitive, metacognitive or socio-affective processes, or a combination of these. The majority of the interventions involved cognitive strategy training, either alone or in combination with training in other strategy types.

Table 3.5 Intervention type studied (N = 38 studies)

Intervention type	N
Cognitive strategy training only	24
Metacognitive strategy training only	8
Socio-affective strategy training only	0
Mixed strategy training (metacognitive and cognitive)	4
Mixed strategy training (metacognitive, cognitive and socioaffective)	2
Total	38

Note: Intervention type categories are mutually exclusive.

3.2.6 Strategies tested

Table 3.6 presents a list of the different strategies used in training. Only the principal strategies have been listed from each study. A total of 35 different interventions or elements of interventions were identified. The number of different types of strategies involved in training has not been presented as it bears little relation to the number of studies; more than one type of strategy was generally involved in any one strategy training intervention.

Table 3.6 Strategies tested

Cognitive interventions
Asking higher order questions
Awareness raising
Clarifying and checking
Contextualisation
Delaying speaking
Dictionary strategies
Focusing on specifics/selective attention items
Grouping/recombination
Identifying task purpose
Ignore and continue
Inference
Input- and output-based instruction
Inventories
Keyword, mnemonics and association strategies
Learner diary
Notetaking
Predicting
Referential links (discourse/grammar)
Reflection and guided reflection
Revision and redrafting
Semantic mapping, glossing, précising
Summarising
Text structure and content exposition
Metacognitive
Awareness raising
Planning
Self-evaluation
Self-monitoring
Verbalisation
Affective
Avoiding frustration
Co-operating with peers
Deep breathing
Not giving up
Rehearsal
Self-rewarding
Self-encouragement

In many cases, the study interventions had multiple facets (and the composition of the intervention may be central to any effect it has), so this table should only be seen as illustrative of the type of interventions found rather than as a comprehensive description of more complex interventions. For example, in the case of input- and output-based instruction, the input-based element of the intervention was measured against an outcome of *comprehension*, while the output-based instruction was measured with reference to *production* of language items.

3.2.7 Intervention length

It was difficult to summarise the different configurations of the interventions due to the variety in their composition and descriptions incorporating numbers of sessions over weeks and months in quantities of hours or minutes, sometimes integrated and other times in discrete amounts, sometimes given as self-access or independent study and sometimes as whole group work. As an illustration of the types found, Table 3.7 shows the diversity. The number of studies that do not report or do not make clear the length of intervention is of concern (N = 7).

Table 3.7 Intervention length (N = 38 studies)

Intervention length	N
2 sessions (90 mins)	1
3 sessions	1
< 2 hours	4
2-5 hours	3
6-10 hours	4
11 to 20 hours	4
>20 hours	4
1-5 weeks	1
6-10 weeks	4
11-20 weeks	3
21-26 weeks	1
27-52 weeks	1
Unclear	7
Total	38

Note: Intervention length categories are mutually exclusive.

3.2.8 Language being learned

As Table 3.8 illustrates, in the majority of studies, the participants were learning English (N=24), either as a second language or foreign language, and while this reflects the high amount of research carried out in ELT in general, just under half of these were of learners of English as a foreign language similar to learners of the other foreign languages listed.

Table 3.8 Language being learned (N = 38 studies)

Language	N
English as a second language (ESL)	14
English as a foreign language (EFL)	10
Spanish*	5
French*	4
Japanese	2
German	1
Italian	1
Latin	1
Russian	1

* One study is coded twice as it investigated both French and Spanish language learners.

3.2.9 Outcomes measured

A wide range of outcomes was measured in the different studies and, as might be expected, there were often more than one or two outcomes per study. Table 3.9 has grouped these into 20 different types of outcome. In some studies, attempts were made to triangulate the effect of the strategy training with different measures, some global proficiency, some more narrowly focused. Over the mapped studies, the outcomes measured in order to demonstrate the effect of the strategy training were as shown, although sometimes there appears little to choose between what appears to be different expressions of the same thing. This is not a comprehensive picture as it presents only the principal outcomes measured in any study for the purpose of giving a broad idea of the range of major outcomes; see Appendix 3.1 for a breakdown of the studies by outcomes measured.

Table 3.9 Outcomes measured (N = 38 studies)

Outcome	N
Comprehension	19
Strategy use	10
Global proficiency	8
Vocabulary acquisition	6
Grammar accuracy	5
Awareness	3
Accuracy of output	2
Aural ability	2
Oral ability	2
Recall (content/meaning units)	2
Self-perception of ability	2
Writing ability (accuracy and quality)	2
Asking and answering higher order questions	1
Attitude	1
Strategy transfer	1
Interpreting meaning	1
Metacognitive knowledge	1
Strategy awareness	1
Unclear	1
No measurement of outcome	1

Note: Studies could look at more than one outcome.

Five studies had some form of follow-up testing, but none of these was any longer than a month after the end of the intervention and none retested exactly the same sample.

3.2.10 Education sector

As illustrated by Table 3.10, the majority of studies were in the adult, tertiary or higher education sector (N=29; 73%). Of those conducted in middle and secondary schools (N=11), all but two were studies whose participants were learning English (N=9). None of the studies was conducted in primary schools.

Only one study set in schools was found that looked at strategy training in languages other than English – and the language in question was Latin. While, strictly speaking, Latin is not exactly a modern language, it was felt that there was no reason to exclude it.

Table 3.2.10 Education sector (N = 38 studies)

Education sector	N
University	20
Secondary school	11
Adult	6
Higher education (non-university)	3

Note: Two of the studies ranged across different education sectors.

See Appendix 3.2 for a breakdown of the studies by education sector.

3.2.11 Country studied

Almost half (N=16, 42%) of the mapped studies in strategy training were conducted in the USA (see Table 3.11). Only four were conducted in Europe, of which only one was in the UK. The rest are spread around the world in Asia (7), Canada (4), Australasia (4) and North Africa (1). Of the randomised controlled trials (RCTs), nine were carried out in the USA and six elsewhere. Of the studies carried out in the USA, all were controlled or randomised controlled trials except for one case study and one action research study.

Table 3.11 Country of study (N = 38 studies)

Country	N
USA	16
Canada	4
Australia	3
Hong Kong (China)	3
Japan	2
Egypt	1
Germany*	1
Korea	1
The Netherlands	1
New Zealand	1
Poland*	1
Singapore	1
Turkey	1
UK	1
unclear	1
Total	38

Note: No studies took place in more than one country.

(*Country deduced as it was not stated explicitly in the reports.)

3.2.12 Studies by publication type and date

With an increasing number of research reports both at master and doctoral level being added to electronic indexes, and therefore more readily amenable to

searching outside libraries, a high number of doctoral theses were found while searching for this review on strategy training research (see Table 3.12).

Just under half the studies found (18 out of 38) were not published in a journal, so a degree of publication bias still exists in as much as the standard literature in the field tends *not* to cite unpublished or postgraduate research. This is exacerbated by the fact that, since 1995, almost the same amount of research was published in journal articles and non-journal articles (N=14, N=12), and these together comprise around 70% of the total research evidence available since 1981.

It seems reasonable to assume that over time, with increasing electronic indexing and searching (not to mention systematic reviewing), it is likely that this bias will reduce. It is unknown what influence the journal type has on any bias, although this review sought published and unpublished material in non-English language journals as well as the high impact journals.

Table 3.2.12 Study by publication type and date (N = 38 studies)

	Journal article	Dissertation	Report	Book chapter	Conference paper
1981		1			
1985	1				
1989	1				
1991	1				
1992	1	1			
1993	1	2			
1994		1			
1995	1	1			
1996	1	1	1	2	
1997	3	2			
1998	1	1			
1999	3	1	1		1
2000	3	2			
2001	2				
2002	1				
Total	20	13	2	2	1

3.3 Identifying and describing studies: quality-assurance results

Thirty-eight reports of experimental studies relevant to addressing the review question were found and have been included in the descriptive map. While there is considerable variety in topic and method between the studies, some important patterns can be seen.

In terms of study focus, the map shows that English language teaching and learning (both EFL and ESL) have been studied in a strategy training context more than teaching and learning in other languages. Only 11 of the studies look at

strategy training in the school sector and the majority (29 out of 38) are studies in the higher education sector (university and adult). The USA accounts for a large number of the studies carried out (16 out of 38, or 42%) and the rest are spread around the world in Europe (5), Asia (7), Canada (4), Australasia (4) and North Africa. Study has been focused on a relatively small number of key learning domains. A substantial number of the studies in the descriptive map looked at reading (N=12), vocabulary (N=6) and/or listening (N=6) and together these outcomes accounted for two thirds of the studies found and included in the mapping (24 out of 38).

None of the studies looked specifically at socioaffective strategy training, although several studies do report participants' perceptions, attitudes and feelings as aspects of the training. Most of the interventions were either cognitive strategy training or mixed in a package with metacognitive strategies, and more than half of the cognitive strategy training interventions comprised clarifying and checking as a skill. *In toto* (cognitive, metacognitive and socioaffective), however, 35 different interventions or elements of interventions were identified – some individually composed and others as parts of combination packages – and they were measured on 20 different major outcomes. Comprehension, understandably to some extent perhaps, was the most common (in 19 of the studies). This raises the issue of whether a useful distinction could be drawn between active and passive strategy training.

The length of interventions varied considerably, ranging from 'single deliveries' or events of an hour or less to interventions spread out over a year. A number of the studies (15 out of 38) evaluated interventions of between 2 and 20 hours in length – perhaps a distinction is warranted between strategy training for awareness purposes (which could include short, sharp-shock-type effects) and training that may depend on a cumulative effect over time (practice-based strategies). Whatever the case, there is a clear lack of measurement of the effect over time in *all* the studies; none carried out any *long*-term follow-up measurement that looked at the duration of the effect of training and this would severely limit any evaluation of effectiveness in terms of cost-benefit for example. Five studies included some form of follow-up testing, but none was any longer than a month after the end of the intervention, or retested exactly the same sample.

Concerning research type, 28 of the 38 mapped studies were comparative studies (one or more intervention compared, sometimes with a control group or non-intervention group) and 57% of these were RCTs, where participants from a homogeneous cohort are allocated randomly to one or other intervention (ideally with neither they nor the assessors being aware of who is receiving what).

Finally, the map illustrates the development of research in strategy training for language learning and teaching over time. Coverage in published and unpublished research since 1980 is spread over those years, but there has been a marked increase in the number of studies planned and carried out since the mid-1990s. This may well have occurred as a result of the greater importance, from around that time onwards, given to issues of learner autonomy, self-access learning, independent study and learner training particularly in the English language teaching (ELT) field. However, it would also fit with language providers increasingly needing to rationalise delivery costs and find more effective ways of teaching large, although dwindling, numbers of learners. Dissemination of reports of studies has been primarily through journal articles (52%) and postgraduate dissertations (34%) since the 1980s and the rest in reports, book chapters or conference papers. It has not been possible here to estimate how many of the

postgraduate dissertations have gone on to be published in the mainstream academic literature; until recently, postgraduate publications remained 'grey' literature, and however good they might have been, many did not see the light of day. However, with electronic indexing, this bias has been reduced, and dissertations are increasingly available at least in citation reference form online or on electronic databases/indexes.

The review authors note again here that a small number of additional reports have been identified as potentially relevant for the map during this review but, as a result of time constraints, these reports have not yet been considered further. The status of a further three reports is tentative due to a lack of response from authors. These reports have been identified in Chapter 6 of this review for the sake of transparency but will be considered further when the review is updated.

4. IN-DEPTH REVIEW: RESULTS

This chapter focuses on the studies identified by this review's searches and considered to be the most appropriate to answer the review question, 'What is the effectiveness of strategy training?'. After setting the 25 studies found in the context of others found within the literature, each study is described in some detail. The characteristics and findings of all 25 are then considered as a whole, in terms of how much they help answer the review question.

4.1 Selecting studies for the in-depth review

Application of the criteria for the in-depth review (see section 2.3.1) resulted in 25 of the 38 mapped studies being included for in-depth evaluation. These are listed in full in section 6.1. The studies excluded at this point, with reasons for exclusion (N=13) are listed in section 6.2.

The 13 excluded studies include ten studies that did not use an experimental study design. Also included were three that appeared to meet the inclusion criteria for the in-depth review but were retrieved too late for in-depth study in this first version of the review or needed further consideration. These three studies were keyworded and appear in the map, but are not evaluated in-depth for this stage of the review (Baily, 1996; Riley and Harsch, 1999; Stokes, 1981).

4.2 Characteristics of studies included in the in-depth review

All the studies in the in-depth review (N=25) reported a clear strategy training intervention, and incorporated a comparison group in their research design. A total of 15 involved random allocation to comparison groups. The remaining 10 studies, which have been referred to as comparative studies, used a comparison group but did not allocate participants using randomisation. Appendix 4.1 presents a study-by-study description of the 25 studies included in the in-depth review.

Table 4.1 presents the language skills focused upon by the training interventions; in two cases, the strategy training focused equally upon two skills or outcomes. Again, reading was the most studied language domain. There is no apparent particular reason why reading (or reading comprehension) is the most frequently studied while the other skills are addressed in approximately similar proportions; one suggestion is the possibility that it underpins more centrally, as an indicator of learning, the other productive language skills.

Table 4.1 The main language skill focus of training (N = 25 studies)

Language skill/domain	N
Reading	9
Overall language ability	5
Listening	6
Speaking	3
Vocabulary	2
Writing	2

Note: The total is greater than 25 as two studies address two domains (Burgos-Kohler, 1991; O'Malley *et al.*, 1985).

Table 4.2 shows that a minority (N=10) of the studies explicitly categorised the training intervention under study by the type of strategies involved. Those which did so were split evenly between those that looked at metacognitive strategies and those that looked at a mixture of strategies. When all the interventions were classified by reviewers according to the type of strategies involved, the majority were found to be cognitive interventions.

Table 4.2 Strategy training type studied (N = 25 studies)

Strategy training type	N as reported by study authors	N as interpreted in review
Cognitive strategy training only	0	14
Metacognitive strategy training only	5	7
Socio-affective strategy training only	0	0
Mixed strategy training		
Cognitive + metacognitive	3	2
Cognitive, metacognitive + socioaffective	2	2
Not stated/unclear	15	0
Total	25	25

The picture is complex and serves to highlight the questions both of defining the strategies, and of what actually matters; it is probably more of academic interest than real relevance (to learners at least) that a strategy is cognitive, metacognitive or socio-affective. This said the difference is illuminating in as much as it points out at least two features: firstly, the relative absence of research into training interventions involving socio-affective strategies, despite a renewed interest of late in motivation and attitudes of learners, and in intercultural competence; secondly, the research identified for the review has focused more on getting learners to do things (cognitive) rather than to know or notice things only (metacognitive), although the distinction is somewhat simplistic, as the latter does not prevent learners taking action themselves.

A study-by-study account of the strategies involved in these training interventions can be found in section 4.3.

4.3 Further detail of studies included in in-depth review

This section summarises the included studies one by one, with brief descriptions of each intervention and the method used for its study. The reviewers' appraisals are incorporated and findings are summarised in terms of the overall weight of evidence the study provides in addressing the review question. The studies are grouped according to the main skill, or domain targeted by the intervention. Preceding each group of studies is a table that aggregates the findings of these studies. This should permit the reader to get an overall handle on the number of studies and their weights of evidence for each skill and to locate each study within this. Tables detailing key study characteristics, findings and the weight of evidence decisions in full are presented as Appendices 4.1 and 4.2.

4.3.1 Speaking ability (N = 3 studies)

Summary of evidence

Findings	3 positive
Reliability	1 high, 2 medium
Relevance	1 high, 1 medium, 1 low
Weight of evidence	3 medium

Can learner strategy instruction succeed? The case of higher order questions and elaborated responses

Ayaduray and Jacobs (1997) conducted an RCT to evaluate the effect on oral skills of 10 weeks of training of two groups of secondary school learners (N=32) of English in Singapore in asking and answering higher order questions. According to the study's authors, the rationale for this is the importance of the role that question asking plays in the promotion of thinking skills. Following the intervention, the learners were recorded participating in group discussions. In the analyses of the contributions comparing the intervention and non-intervention group, the study found that the intervention group more frequently asked higher order questions and gave elaborated responses. They conclude that their results support the view that it is possible to train students to adopt new, more effective learner strategies: with the appropriate preparation (providing instruction and opportunities to ask higher order questions) and that particularly in this area, students can be trained therefore to become more effective questioners. They go further and propose that such training should be integrated into language instruction.

The small sample (two groups/32 students), the unclear randomisation procedures and the absence of blinding in the allocation and assessment are limitations to the strength of the findings. The intervention group was reminded for the post-testing to use the strategy that had been taught and this may have exerted an influence on their performance beyond the effect of the intervention itself. The reliability and relevance of the study were both considered to be medium, as was its overall weight of evidence in addressing the review question.

The effect of metacognitive strategy training with verbalisation on the oral accuracy of adult second language learners?

Holunga (1994) trained a group of nine participants in a study of 48 adult advanced English learners in Canada to use metacognitive strategies, comprising predicting, planning, monitoring and evaluating. Another group in the 48 were given metacognitive strategy training with verbalisation, while a third group was used as control. The study used an RCT design. On testing of discrete item oral accuracy (verb forms), the metacognitive strategies with verbalisation group improved the most, although the group without verbalisation also improved, while the control group showed no improvement. The differences were significant. One caveat is that the performance of the two intervention groups was on particular tasks that focused on form; there were no measures of general improvement or overall proficiency changes. Delayed post-testing after a month found that the effect of the intervention had lasted.

While well-conducted, this study carries a medium weight of evidence in answering the review question as the generalisability of the findings to any other population than the study sample is very limited – that is, a group of advanced adult ESL learners, all of whom already spoke two other modern languages.

Learner strategy applications with students of English as a second language

O'Malley *et al.* (1985), in an RCT of 75 secondary school 'intermediate' learners of ESL in the USA set out to evaluate the effect of strategy training (metacognitive, cognitive and socioaffective) on learners' speaking, listening and vocabulary skills. In the experiment, a control group carried out the same task but without the strategy training received by the two intervention groups. The training was carried out in 50-minute lessons over eight days, and the outcomes were assessed in listening and speaking tests.

The data showed a tendency towards better performance for the two intervention groups on the listening task, but the association was statistically non-significant. On the speaking task, the intervention group performed better, and the result was statistically significant. The vocabulary intervention and findings are not reported.

No baseline measurements are reported and no data are given for pre-test or interim test performance, and so the conclusions of the study need to be treated cautiously. The study was judged to be of high relevance, but of low reliability and to carry a medium weight of evidence for answering the review's question.

(Note: This study is also reported in section 4.3.6 on listening.)

4.3.2 Reading ability (N = 9 studies)

Summary of evidence

Findings	6 positive, 1 negative, 2 mixed
Reliability	2 high, 3 medium, 4 low
Relevance	7 high, 2 medium
Weight of evidence	2 high, 3 medium, 4 low

Effects of strategy training on reading comprehension in first and foreign language

Bimmel *et al.* (2001), in a comparison study run over 15 weeks, tested a group of 131 Dutch secondary school students (12 in the experimental group and 119 in a control group) to see whether the reading comprehension strategies taught to them were successful, firstly for reading in Dutch, and secondly for reading in English. The reading strategies involved comprised the following:

- identifying key fragments in text
- identifying hinge words (the connectors that give a handle on a text)
- questioning (the learner's interrogation of the text)
- semantic mapping (having an understanding of the main units of the text)

Learners were trained in these in two modes: explanation method and consciousness-raising.

There was no significant difference as measured by comprehension between the groups concerning transfer of the training effect to English. However, the authors found significant differences in favour of the intervention group on the incidence of identifying key fragments, semantic mapping and hinge words (but not on questioning) that provided evidence of the effect of the strategy training for Dutch, (they conclude that the reading strategies training for their first language was effective).

The study is of medium relevance to the review question as the training takes place for the learners' first language, and it is only when applied to their learning of English that it becomes germane to the review. Due to aspects of sample size and selection, and significant differences in variance on two out of seven baseline measures, there are limitations to the study, mainly because the strategy training is for the learners' first language and it is only the transfer of the effect that is assessed in their second language. The authors discuss many of the limitations themselves but the reliability of the study was judged to be low. The researchers also state that they explicitly pressured the participants to use at least one of the strategies in the Dutch and English reading comprehension tests; this may have influenced the effect of the intervention and any autonomous deployment of the strategies by the learners. The overall weight of evidence in addressing the review question was judged to be low.

Metacognitive strategy training for ESL reading?

Carrell *et al.* (1989) used a comparison design to evaluate the effect of two metacognitive interventions on reading ability on 18 of a group of 26 adult, mixed linguistic background university language learners in the USA. The interventions comprised semantic mapping (explicitly organising and categorising information before reading a text, then 'mapping' it against the text) in one group and experience text relationship (ETR) in the other (training the reader to activate personal knowledge in order to improve comprehension of the text).

The authors found no significant difference between the intervention groups and comparison groups on comprehension as measured by multiple-choice questions. There were significant differences in favour of the intervention groups on open-

ended questions. ETR scored more highly than semantic mapping on a 'partial semantic map' and ETR scored more highly on open-ended semantic mapping. They did, however, find that learning style and type of intervention had a significant influence on the effect. Overall, the authors claim that metacognitive training, in this case semantic mapping and ETR, enhances reading.

However, the small sample size (four groups and 26 students), and some aspects of the testing of comprehension limit the reliability of the study findings. Specifically this includes the similarity between intervention and testing – semantic mapping was part of the test as well as the intervention; the control group would not be likely to perform well on the test, yet this could not be a valid assessment of reading comprehension ability. Overall, the study was considered to carry a low weight of evidence for answering the review question.

Effects of three semantic mapping strategies on EFL students' reading comprehension

EI-Koumy (1999) in a study similar to that of Carrell *et al.* (1989) investigated three variations of semantic mapping on reading comprehension (English) on a group of 60 students in a total study population of 237 students majoring in French in an Egyptian university. The study used an RCT design. The intervention of interest to this review trained learners to generate their own maps of the texts to be read, and the study author reports that there was no difference between student-mediated and teacher-initiated semantic mapping, but that teacher-student interactive mapping was significantly better than the former two. This is to be expected but we do not know what happens when the scaffolding provided by the teacher is no longer there. The author concludes that more research is needed. Overall the study was considered to carry a medium weight of evidence for answering the review question.

Referential strategy training for second language reading comprehension of Japanese texts

Kitajima (1997), in a study of 28 American college students studying Japanese trained the participants in the intervention group to identify phrasal divisions by listing verbs and conjunctions, to identify logical connections between clauses by specifying functions of conjunctions and adverbials, and also forming questions based on verbs and cohesive devices, and then comparing the analysis with what they found while reading. A control or comparison group had taken the same course a semester earlier, although this is a limitation given that the sample is not from the same frame.

For the outcome of overall reading comprehension, the study reports no significant difference between groups on paragraph 1 of the test text that was used, but significant differences between the groups on paragraphs 2, 3 and 4. On the second outcome, that of identifying referential ties in the texts, no differences were found on paragraphs 1 and 2, while significant differences were found on paragraphs 3 and 4. The author concludes from this that, in spite of the limitations, 'strategy training that directs students' attention to monitoring coreferential ties can be considered to have positive effects on their comprehension of Japanese texts' (p 1).

The author discusses the limitations of the study which include the possibility of influence from confounding variables and the interactivity with variables other than the one of interest. Selection methods of the samples, absence of measurement

of strategy use after the intervention, and the possibility of experimenter bias are also limiting factors that lead to this study carrying a low weight of evidence.

The effect of metacognitive strategy training on reading comprehension and metacognitive knowledge

Kusiak (2001), in a comparative study on 78 out of 158 secondary school learners of English, evaluated the effect of training in metacognitive strategies on reading comprehension and metacognitive knowledge. The intervention consisted of eight 45-minute lessons, additional to their regular language study, in which they were made aware of strategies, practised basic reading strategies, observed their own and peers' use of strategies, and worked on assignments at home. (The experimental group also had a few extra lessons but the teachers were asked not to teach reading skills.)

The study reports positive findings for self-evaluation skills and reading comprehension, and the author concludes that the study points to the effectiveness of metacognitive strategy training for students of intermediate level, and that it was more effective for learners of lower ability. Overall, the study was considered to carry a medium weight of evidence for answering the review question.

Generative précising as a reading comprehension strategy for adult ESL learners

McGuire (1999), in a study of 54 out of 71 fee-paying adult learners on an English for professional purposes course in the USA, evaluated the teaching of a strategy in a short intervention of around an hour's duration. The study used an RCT design. The strategy comprised the comparison of two versions of highlighting meaning in text with generative précising, in which the learner makes notes in short translated chunks in the margin of the reading text. The three interventions were underlining and signalling, re-reading and repetition, and then the generative précising; a 'control' fourth group had a grammar topic but it is unclear to what extent this genuinely provided a control group for comparison purposes as it is not stated if this was standard practice and it could itself have exerted an influence on performance.

While the study found a positive effect of generative précising on reading comprehension, the robustness of the findings is compromised by a number of the study's characteristics, including queries surrounding the choice of interventions for comparison, a significant difference between low ability and high ability learners at baseline, unclear evidence as to what extent the strategy of précising was being used by the three or four groups (control group not reported, and only absence of evidence reported for re-reading group). While the interventions were randomised to the four groups, there is no indication of any concealment of allocation or of how the individuals assigned themselves to the four groups, other than them being constrained by scheduling.

While the author discusses the potential problem of the difference between generative précising (the intervention of interest) and underlining, and states that the former is 'meaning-creation' while the latter is 'selection', the study still compares them as a means of evaluating the effectiveness of précising. In light of this, it seems self-evident that the former would prove relatively effective unless both can be clearly shown to influence reading comprehension; this seems to be borne out by that fact that the re-reading and underlining are apparently not more

effective than the control group activity. Furthermore, the control group gain scores were higher than the underlining and re-reading gain scores: in the higher ability group, the gain score of the control group was very close to that of the précis group. As a result of all of the above, the study was considered to carry a low weight of evidence for answering the review question.

The effect of note-taking strategy instruction on comprehension in ESL texts

Najar (1997), in a randomised controlled trial on 135 of 338 college learners of English in Japan in 10 groups allocated to intervention or control, evaluated the effect on strategy use and reading comprehension of teaching them how to take notes.

The results showed that note-taking had a positive effect on reading comprehension (comprehension test), and that, within the intervention groups, there was a difference between those who used the strategy and those who did not. However, some selection of data to use appears to have occurred as 'only learners who used the note-taking strategy on the post test were used to measure notetaking strategy transfer' (p 97). The effect did not show any greater use on the 'transfer' task (that is, a second similar comprehension test).

No information is given on dropouts or non-completers (at least nine students in the intervention group), and a difference found between intervention groups was ascribable to teacher differences. Overall the study was considered to carry a medium weight of evidence for answering the review question.

The effects of structure strategy training on the recall of expository prose for university students reading French as a second language

Raymond (1993), in a randomised controlled trial of 43 first-year university learners of French in Canada, trained the intervention group in strategies to derive the content of a reading text by unpicking and describing its structure: five frequently found top level structures (TLS) in prose were identified to inform the training: description, collection, causation, problem-solution and comparison (although only problem-solution was tested). The reading component of a proficiency test was given to all participants before the study, and both groups received five hours of training. An outside instructor taught the intervention group covering each TLS in a session, while the control group spent a comparable time reading without training.

One month after the intervention, all students were given a similar test using a second text. The study found that the experimental group scored higher in content recall than the control group at post-testing – but only on one of the texts used. The researchers found that some more proficient students were already using TLS strategies – perhaps spontaneously transferring skills from their first language (see Bimmel *et al.*'s study which did not find evidence of transfer of taught strategies from first language to second language). The author concludes that there is some evidence for a positive effect of strategy training using the TLS strategy, the statement being restrained by discussion of the complexity of strategy interventions and interconnectivity of factors in educational interventions. Interestingly, the author writes that a clearer picture will in the end come from accumulation of evidence from research.

The robustness may be limited by the cluster randomisation (not individuals) in conjunction with the small sample size. Students were paid to take part in the study, and it was not reported whether or not allocation was concealed. However, overall the study was considered to carry a high weight of evidence for answering the review question.

Metacognitive strategy training for reading: developing second language learners' awareness of expository text patterns

Talbot (1995), in a randomised cluster trial (12 intact groups rather than students randomised to intervention and control arms) of 183 + 51 in a total of 244 Chinese background Hong Kong learners from post-secondary training colleges, trained the intervention group in metacognitive awareness of text structure in English. The control group continued with its standard syllabus, and both arms were pre- and post-tested at five weeks; the intervention group was tested again after four months to assess the duration of any training effect – this is one of the few studies that incorporate delayed post-testing.

The primary outcome of interest is performance on reading comprehension tests, but the effect of the intervention was assessed against other variables including gender, L2 proficiency level, self-rating of reading ability, and learning approach.

The study found that the intervention groups made statistically significant gains over the control groups with gains on three out of the four components of the testing, but not on the component using textual clues to reconstruct scrambled expository text. The group classified as 'medium proficiency' gained more than the other two levels (low and high) and this difference was significant. Removing the controls from this analysis, the low and medium both gain more than the high level learners. Qualitative data were also favourable to the intervention groups. The author concludes that strategy training in expository text structure awareness was effective in its influence on learning performance. Delayed post-testing after four months relied upon data sampled from selected intervention participants and did not involve retesting comprehension. This study was judged to have a high weight of evidence for answering the review question.

4.3.3 Writing ability (N = 2 studies)

Summary of evidence

Findings	2 positive
Reliability	2 medium
Relevance	2 high
Weight of evidence	2 medium

Using quality and accuracy ratings to quantify the value added of a dictionary skills training course

Bishop (2001) evaluated the effect of training a group of 15 out of 30 adult distance learners of French to use a dictionary when redrafting and revising an essay. The training consisted of learners spending three to six hours reading and working through the dictionary course that they received, and then redrafting an essay that they had written. A control group, although not from exactly the same

body of students, redrafted their essay without having the dictionary skills course. Bishop found that the intervention group improved by around 14% and the control group by about 1.5% in language accuracy scores on the redrafted essay, and 11% and 2.2 respectively for the two groups on quality scores at redraft.

The study is highly relevant to the question of the review regarding strategy training. While there are limitations to the reliability of the findings of the study due to aspects of the sample selection, absence of measurement of strategy use (how much and in what ways did they use the dictionaries?) and analyses of variance in the results, the simple intervention with clear outcome measurement and the apparently large effect size of >10% improvement make this a noteworthy study. The study was judged to be medium weight of evidence in answering the review question.

An investigation into the effects of revision strategy instruction on L2 secondary school students

Sengupta (2000), in a study carried out around 1997, although not published until 2000, evaluated the effect of getting secondary school learners in Hong Kong to redraft their essay first drafts. The strategies for redraft included training in making student texts more reader-friendly by unpicking variables such as attitude to writing, and student definition of a good composition, while the main outcome measured was gain score on the rewritten essay. Six compositions were treated during the year and teacher scaffolding was gradually reduced. The study population of 118 girls, in pre-existing class groups, was allocated to two intervention groups and two control groups (although one of the latter was excluded before the intervention began); details of selection are not reported. This study was classified as having a comparative study design.

The compositions were marked holistically, and gain scores compared after redrafting. Students' perceptions were recorded by questionnaire (only the intervention groups), and a sample (N=8) of participants was interviewed towards the end of the study. An interesting aspect of the study was that it preserved as much as possible of the standard school situation

The study reports that the two intervention groups made greater gain scores than the control groups, and concludes that the findings corroborate the theoretical belief that revision has the potential of a new assignment, and thus may be worth the time and effort. However, analysis did not control for what happened in the classroom: that is, differences between classes, teachers and delivery of intervention. Data are reported on 100 students only, although 118 took part, and the sample was female only and this cannot be controlled for. This study was judged to have a medium weight of evidence for answering the review question.

4.3.4 Overall language ability (N = 5 studies)

Summary of evidence

Findings	2 positive, 1 negative, 2 mixed
Reliability	3 med, 2 low
Relevance	3 high, 1 medium, 1 low
Weight of evidence	3 medium, 2 low

The effect of a selected group of language learning strategies upon language development

Burgos-Kohler (1991) used an RCT design to examine whether students learning Spanish in an American university, trained in keyword, elaboration, association, grouping, recombination and contextualisation strategies for vocabulary learning using an embedded instruction approach would improve their Spanish vocabulary and their overall proficiency in Spanish more than those not receiving training for Spanish vocabulary.

This six week study compared the achievement of students in three groups. The students in experimental group 1 were given instruction on various language learning strategies and were provided exercises in which to apply these strategies. Students in experimental group 2 were only given lists of vocabulary words to study and use in sentences. Students in the Control Group were left to their own learning devices. (p vii)

Statistically significant positive results were reported for both vocabulary acquisition measures and end of semester grades. Confounding factors, including the effect of the intervention group being given additional materials to work with as self-study, selection of the test vocabulary items from the beginners course books, possible unclear division between intervention and test, may affect the reliability of the study, although the study is highly relevant. The study is therefore considered to carry a medium weight in addressing this review question.

(Note: This study is also included in section 4.3.5 on vocabulary.)

Explicit instruction in grammar: a comparison of input-based and output-based instruction in second-language acquisition

Cadierno-Lopez (1992) compared the impact of two different forms of instruction and the way learners process incoming linguistic information. The study used an RCT design and involved six groups of Spanish learners at an American university. The study investigated teaching learners to focus

- explicitly on comprehension of grammar *input*, or
- explicitly on producing examples (*output*) of the grammar item compared with a group that received no training.

The researcher hypothesised that the former is more effective than either the latter or no training on either vocabulary acquisition or overall proficiency, both in terms of understanding the input and in accuracy of output.

There was no significant difference between processing instruction and traditional instruction, although both were statistically significantly better than no instruction at all. However, the researchers discuss the possibility that the repeated testing could have had an instructional effect itself and contributed to the effects found.

The study is not highly relevant to the review question as it is to some extent an investigation of two teaching methods. However, the awareness aspect embodied in it brings it into the realm of strategy training from the students' point of view. The study was considered to carry a medium weight in addressing this review question.

Consciousness raising and strategy use

Feyten *et al.* (1999), in the second phase of an RCT with 863 secondary, middle school and university learners of French and Spanish, in the USA, looked at the effect of giving a metacognitive awareness raising (MAR) handout on 26 strategies, and a cognitive awareness raising (CAR) handout on 26 reasons for studying a language. These were compared with a control/placebo group that completed a survey on myths about learning languages, although no information was available in the report on the relative numbers of experimental and control groups. The effect of these 'short sharp shock' awareness-raising interventions was assessed via the participants' performance on their regular final examination grades.

The findings were mixed, including a non-significant positive effect in the control group amongst the university learners of French, a non-significant positive effect of CAR in the secondary school group, and a greater effect in the control group of University Spanish and French learners. The findings in support of CAR were non-significant and only in one group (secondary school), yet the authors conclude that MAR and CAR seemed to be having some effect on learners. The study was considered to carry a low weight in addressing this review question.

A two-phase study involving consciousness-raising and strategy use for foreign language learners

Flaitz and Feyten (1996) in phase 1 of their two-phase study, a randomised trial, looked at the effect of strategy awareness raising on 130 (intervention group) 99 (control group) first-year university learners of Spanish. The intervention comprised a handout called 'How to survive Spanish 1 or 2', outlining 26 language learning strategies and an activity to categorise these strategies: the control groups received a 'placebo questionnaire' but this may have had some effect on the learners, if it differed from standard practice. The effect was investigated via regular end of semester tests and questionnaires about strategy use.

The findings included a significant difference between experimental and control groups in favour of the intervention as measured by final grades, although the authors also discuss the possibility of confounding variables (such as teacher differences, enthusiasm) playing a part, and they conclude from their results that awareness-raising in a short sharp burst has the potential to be effective. The study was considered to carry a medium weight in addressing this review question.

Language learning strategies advice: a study on the effects of online messaging

Meskill (1991), in a study with a group of 34 ESL learners at a university in the USA, looked at the effect of strategy advice messages appearing on screen as they worked on video language tasks online. An RCT design was used and one group received messages while the other did not. Participants were observed to ascertain whether or not they read the messages, and whether or not they followed the advice in them. Outcomes of interest measured in the study by observation were whether learners stayed on task longer, read the messages, took the advice and performed as well as good language learning strategists on a post-test of oral ability and attitude.

While the quantitative findings show that the messages have no significant effect on performance, the author still concludes from the qualitative data that student reactions to the messages are suggestive of positive effects.

However, insufficient data in the report, inconsistencies between the reported results and conclusions, the absence of pre-testing, and sparse information on participants make it difficult to see to what extent the performance and intervention were affected by participant ability and characteristics, or other influences. The study was considered to carry a low weight in addressing this review question.

4.3.5 Vocabulary ability (N = 2 studies)

Summary of evidence

Findings	2 positive
Reliability	2 medium
Relevance	2 high
Weight of evidence	2 medium

The effect of a selected group of language learning strategies upon language development

Burgos-Kohler (1991) used an RCT design to examine whether students learning Spanish in an American university, trained in keyword, elaboration, association, grouping, recombination and contextualisation strategies for vocabulary learning using an embedded instruction approach would improve their Spanish vocabulary and their overall proficiency in Spanish more than those not receiving training for Spanish vocabulary.

This six week study compared the achievement of students in three groups. The students in experimental group 1 were given instruction on various language learning strategies and were provided exercises in which to apply these strategies. Students in experimental group 2 were only given lists of vocabulary words to study and use in sentences. Students in the Control Group were left to their own learning devices. (p vii)

Statistically significant positive results were reported for both vocabulary acquisition measures and end of semester grades. Confounding factors – including the effect of the intervention group being given additional materials to work with as self-study, selection of the test vocabulary items from the beginners course books, possible unclear division between intervention and test – may affect the reliability of the study, although the study is highly relevant. The study was therefore considered to carry a medium weight in addressing this review question.

(Note: This study is also included in section 4.3.4 on overall language ability.)

Learning and recall of foreign language vocabulary: effects of a keyword strategy for immediate and delayed recall

Lawson and Hogben (1998) trained two intervention groups and a control group in a sample of 40 secondary school girls learning Italian in Australia. The study used

a keyword and an elaborated keyword method to test the effect on vocabulary recall. This study was classified as a comparison study. The intervention (N=26) comprised the participants devising their own keywords for a set of nine nouns in a booklet given to them. The keyword method was then explained to them and how they should use it to remember vocabulary items, and they were encouraged not to abandon the search for keywords even when they found it hard. The control group had no specific instruction on vocabulary learning methods. Over the 10 days after the intervention, the students were tested three times for their ability to recall the listed words.

The study reports a significant positive effect for the elaborated keyword method and the keyword method, but its reliability is limited due to lack of reporting on sampling and sample selection, and the shortage of detail on testing and assessment in the study. Perhaps more importantly, it is difficult to determine the material difference between keyword and elaborated keyword interventions such that they could give rise to significant differences. In such a bundle of strategies as the elaborated keyword method, it is important to unpick the compound effect. The study was considered to carry a medium weight in addressing the review question.

4.3.6 Listening ability (N = 6 studies)

Summary of evidence

Findings	4 positive, 1 negative, 1 mixed
Reliability	3 medium, 3 low
Relevance	1 high, 5 medium
Weight of evidence	3 medium, 3 low

The effect of listening comprehension strategy training with advanced level ESL students

While this comparison study by McGruddy (1995) reports a statistically significant difference in favour of the intervention group, it is only on the outcome of selective attention that the intervention group increased their strategy use. The researcher found that the listening logs were not as productive as anticipated and that prediction and inferring strategies were not frequently mentioned. Overall, the author concludes that training selective attention as a strategy may be useful in improving listening comprehension ability. However, there appears to be a difference between the abstract and discussion sections of the report: whereas the former reports positive intervention group change on the Michigan test for listening, the latter reports no change on this test.

Confounding variables are not discussed in the study and it is possible that bias was introduced by the selection of a specific class for the intervention group. Differential attrition between the groups (failure to complete or do the intervention or remain in the study) cannot be assessed as the attrition is not reported and the numbers are unclear. The study was considered to carry a low weight in addressing the review question.

Learner strategy applications with students of English as a second language

O'Malley *et al.* (1985), in a randomised controlled trial of 75 secondary school 'intermediate' learners of ESL in the USA, set out to evaluate the effect of strategy training (metacognitive, cognitive and socioaffective) on learners' speaking, listening and vocabulary skills. In the experiment, a control group carried out the same task but without the strategy training received by the two intervention groups. The training was carried out in 50-minute lessons over eight days, and the outcomes were assessed in listening and speaking tests.

The data showed a tendency towards better performance for the two intervention groups on the listening task, but the association was statistically non-significant. On the speaking task, the intervention group performed better and the result was statistically significant. The vocabulary intervention and findings are not reported.

No baseline measurements are reported, and no data are given for pre-test or interim test performance, and so the conclusions of the study need to be treated cautiously. The study was judged to be of high relevance, of medium reliability in its own terms and to carry a medium weight of evidence for answering the review's question.

(Note: This study is also reported in section 4.3.1 on speaking.)

Listening strategy instruction for female EFL college students in Japan

Ozeki (2000), in an unpublished doctoral dissertation, reports on the effects of strategy instruction for a group of 25 in 45 female EFL first-year college students in Japan on their listening ability. The intervention was 'embedded, integrated' training that consisted of metacognitive strategies (directed attention, selective attention, self-evaluation), cognitive strategies (note-taking, inferencing, summarisation) and socioaffective strategies (questioning for clarification, co-operation) and the intervention group was compared with a control group that did not receive strategy training.

The effect of the intervention was measured using a comparison study design, via listening comprehension ability, attitude towards the strategies, and the students' use of them. The control group in fact used the cognitive and socioaffective strategies more than the intervention group, although the intervention group used metacognitive strategies more. Overall, the intervention group's use of strategies was higher.

On the outcome of listening comprehension ability, there was no significant difference between experimental and control groups, both of which gained significantly at post-testing. The gain of the control group was greater than half of the gain of the experimental group.

Interestingly, the author concludes all the same that the strategy training was successful (for both groups), that students had positive attitudes towards the strategies (only the experimental group reported on these in journals), and that they used them and continued to use them beyond the end of the intervention (although there were no delayed post-tests). The study was considered to carry a low weight in addressing the review question.

The effects of strategy training on the aural comprehension of L2 adult learners at the high beginning/low intermediate proficiency level

Paulauskas (1994) in a doctoral dissertation for a Canadian university, used an RCT design to look at the effect of training 44 adult ESL learners at beginner/intermediate level in four comprehension-fostering strategies – predicting text content, summarising main ideas, questioning for comprehension, and clarifying comprehension difficulties. One of the intervention groups received reciprocal training (taking the role of experimenter or teacher in instructing the strategies) as well as the strategies, the second group had no reciprocal training, and the control group received the materials with no training in how to use them.

The outcomes were tested via an L2 listening comprehension test, and a specifically designed strategy test, and the study found that the two strategy groups performed better than the control group on the listening tests, but that there was no difference between the two intervention groups.

Due to timetable constraints, some of the participants (N=5) were not allocated at random. In conjunction with small group sizes, this may have affected the robustness of the findings. The study was judged to carry a medium weight of evidence for answering the review's question.

Intervening in tertiary students' strategic listening in Japanese as a foreign language

Seo (2000) conducted a doctoral study, in which 10 university level learners of Japanese in Australia were given cognitive and metacognitive strategy training and the effect on their listening comprehension ability was tested using video news broadcasts in Japanese. The intervention group received training in identifying key terms; elaborating and inferencing, which the author explains are derived from metacognitive strategies of planning; defining goals, monitoring and evaluation; and cognitive strategies of predicting content, listening to the known, listening for redundancy, listening to tone of voice and intonation, and resourcing.

Participants decided, based on their timetable needs, which group to join. There was no concealment of allocation, and learner variability was not discussed. The author discusses limitations of findings due to sample size: the sampling frame was 40 but voluntary participants numbered 10. There was baseline testing of Japanese ability (audio only) and the post testing was audio-visual.

The study reports positive findings, that the intervention group performance overtook that of the control group in the two final tests in a series of eight (tests were conducted on both groups from the outset each week and a possible effect from the test must be considered – as does the author – as well as the performance on the first six tests. The study was judged to carry a low weight of evidence for answering the review's question.

Can strategy instruction improve listening comprehension?

Thompson and Rubin (1996) conducted a study involving 36 third-year Russian learners in an American university. They used an RCT design to evaluate the effect on their listening comprehension of listening strategies incorporated into 45 video clips, amounting to 15 hours, watched over the academic year; the control group had the same clips. The intervention and control groups had different lesson plans.

The training consisted of metacognitive strategies of planning, defining goals, monitoring, evaluating; cognitive strategies included predicting content, listening to the known, listening for redundancy, listening to tone of voice and intonation, and resourcing (for example, jotting down phrases to see what they mean).

At the end of two semesters, both groups were tested using the same video and audio comprehension tests. The following year, the group taking third-year Russian course was given the same intervention and test, and their results were added to those of the intervention group of the previous year.

Authors accept that the sample is small and lacks power. In addition to this, there are risks to reliability from the use of historical data for control group comparison, and the combining of data from two years – indeed without a control group, data for the third intervention group. The study was judged to carry a medium weight of evidence for answering the review's question.

4.4 Synthesis of evidence

This section brings together the findings as reported in the previous section in detail for each study. The aim is to provide a narrative overview of the findings for each domain covered: speaking, reading, writing, overall language ability, vocabulary, and listening.

The synthesis for each domain refers to the findings as reported by the study authors, the most salient features of the studies, and the weight of evidence evaluation arrived at in the review in relation to the questions asked.

The main question asked by the review (see section 1.6) concerned the evidence of effectiveness for strategy training, and this remains uppermost. As no studies were found on the domain of motivation, this is not covered here.

The second question posed by the review is not addressed as strategy training *does* work according to the research evidence and therefore to speculate why it does not work is no longer relevant.

4.4.1 Speaking ability

The three studies of strategy training to improve speaking ability are of mixed relevance and reliability with regard to the review question, but the review finds that there is some reliable evidence in favour of a beneficial effect of strategy training on speaking ability. The evidence is not very strong due to small samples, highly-specific population (Holunga, 1994), unclear randomisation procedures, absence of concealment of allocation in the studies and inadequate reporting of data. All three studies report positive findings, but while there is a consensual picture concerning the effect of training on the major outcome of speaking ability, the differences in the interventions and between the studies must be borne in mind. The studies were from three different countries, although all were of ESL learners and this may further constrain generalisability to other languages and other settings. Two of the studies looked at school level learners while the Holunga study participants were advanced adult, multilingual learners.

Authors (date)	Strategy training type	Results as reported
Ayaduray and Jacobs (1997)	Training in higher order questions and elaborated responses	Better questioning, better group discussions
Holunga (1994)	Metacognitive training (predicting, planning, monitoring and evaluating) with and without verbalisation	Greater oral accuracy in both intervention groups
O'Malley <i>et al.</i> (1985)	Mixture of cognitive, metacognitive and affective strategy training on listening, speaking and vocabulary (although the latter is not reported)	Better speaking

4.4.2 Reading ability

Compared with studies looking at other outcomes of interest, those testing interventions on reading comprehension show fewer differences in the approaches taken, both in the interventions (mainly metacognitive/semantic mapping) and in the outcomes of interest (mainly reading comprehension). This makes synthesis of the findings less problematic. Seven of the nine studies on strategy training for reading are considered of high relevance in addressing the review question; most of them look at reading comprehension as the outcome, and most are interventions of semantic mapping or structural mapping to improve comprehension. More than half the studies (five out of nine) are either of medium or high reliability and the overall contribution in terms of weight of evidence of these studies to addressing the review question is quite high (six out of nine are high or medium). In summary then, a body of reliable evidence emerges to show that strategy training for reading comprehension is effective.

Two of the studies found a differential effect depending on the ability level of learners, and this may indicate an avenue for further research to establish where the strategy training might be more effectively focused if it is not a general benefit. A third study found similar differences for level, but the differences were already present at baseline.

The studies were carried out between 1989 and 2001, with most being completed in the late 1990s in the Netherlands, USA, Egypt, Poland, Japan, Hong Kong and Canada; the languages involved were mainly English, but also Japanese and French. This diversity of setting plus the generally cautiously understated positive findings may support a degree of generalisability.

Sample sizes varied considerably (between 28 and 338) and none of the studies randomised individuals – a characteristic that has a bearing on the robustness of any findings. With cluster analysis (that is, randomisation of groups), the power to detect an effect is reduced.

Other aspects of some of the study designs and methods used that constrain the reliability include the following:

- small sample sizes and potential selection bias
- influence of experimenter on the study
- absence of randomisation (only five of the nine were RCTs)
- variance (heterogeneity in groups at baseline)
- validity issues on assessment of reading comprehension

- confounding variables and their non-investigation
- lack of information of dropouts and non-completers
- absence of concealment
- overstated reporting of results
- poor reporting

On the other hand, the Talbot study was one of the very few of the included studies to incorporate any longer-term, follow-up measurement of the intervention effect (delayed post-testing), although it was qualitative and of selected participants and did not replicate the quantitative immediate post-tests.

Authors (date)	Strategy training type	Results as reported
Bimmel <i>et al.</i> (2001)	Do reading strategies (semantic mapping inter alia) taught in first language (Dutch) transfer to English?	No transfer, even though the strategies worked in first language (but some indication of partial transfer)
Carrell <i>et al.</i> (1989)	Metacognitive training (semantic mapping) for reading comprehension	No effect on comprehension as measured in multiple choice questions (MCQs) but a positive effect when measured by open-ended questions
EI-Koumy (1999)	Semantic mapping for comprehension	Positive effect found when accompanied by high degree of scaffolding from teacher
Kitajima (1997)	Identifying certain text discourse features to improve comprehension	Some indication of a positive effect (but perhaps overstated)
Kusiak (2001)	Metacognitive strategy training – practice of reading strategies, observing peers (working on assignments at home) for metacognitive knowledge and reading comprehension	Positive findings, and more effective for low ability learners
McGuire (1999)	Notetaking (generative précising – translated notes of quantities of meaning in the margin) to help reading comprehension	Positive findings with generative précising helping reading comprehension
Najar (1997)	Notetaking for reading comprehension	Mixed findings, although reported as positive effect
Raymond (1993)	Understanding text structure to help in recall of contents	Some indication of positive effect, although author includes caveats on this
Talbot (1995)	Metacognitive strategy training (awareness of text structure)	Positive effect, and subgroup classified as ‘medium’ proficiency’ doing better than high and low on reading comprehension tests

4.4.3 Writing ability

It is difficult to make claims for strong evidence on the basis of two studies that show strategy training in writing has a positive effect, and one might even claim it is intuitive common sense that revising and rewriting of first drafts of essays is likely to produce better written work. However, in both the studies above, the control groups did not show the same improvement, although they also redrafted their work, and one conclusion at least is that the strategy training had something to do with this. In the Bishop (2001) study, it was at the very least redrafting plus use of the dictionary that made the difference, and in the Sengupta (2000) study, carried out over a school year in a relatively naturalistic setting (that is, the intervention was integrated with normal practice), it was redrafting plus removal of teacher support/scaffolding nurturing learners towards greater independence. Both studies were considered highly relevant to addressing the review question, and both were of medium reliability and weight of evidence.

Differences between the studies impinge on the degree to which one can ‘lump’ the findings where one was an adult distance learning population in the UK taking French, the other a female school population learning English in Hong Kong. Generalisability is likewise constrained by the limited evidence available.

Authors (date)	Strategy training type	Results as reported
Bishop (2001)	Training in dictionary skills	>10% improvement of quality and accuracy in redrafted essays
Sengupta (2000)	Redrafting of essays	Positive effect on writing

4.4.4 Overall language ability

Four out of the five available studies were of high or medium relevance to the review question. (The online messaging study by Meskill was less directly relevant as it was less conceived as a strategy training than as a learning method).

The findings of the studies lend some support to a possible beneficial effect of strategy training on overall language ability, although it is not incontrovertible by any means. A characteristic of the approach in two of the studies was that they retained a naturalistic element and measured the effect of the training against the regular end-of-term tests – although this perhaps rendered detection of the effect more problematic, particularly if it is small. The findings are mixed, and the studies discuss both the methodological reasons and the confounding influences that might have played a part in the results.

The confounding factors include the following:

- additional self-study materials given to the intervention group (but not the control group) in one study
- unclear separation between the test vocabulary items and the content of the course book (so any effect could not be clearly attributed to strategy training rather than familiarity through the materials)
- repeated testing possibly having an instructional effect in one study

- teacher differences, including enthusiasm, which affect delivery of the intervention)

Methodological considerations that affect the reliability include the following:

- non-randomised selection of samples (although 3 of the studies were randomised)
- insufficient background data and reporting of characteristic of the participants
- no long-term follow-up of the duration of any effect
- little measurement of deployment or use of strategies by learners

The findings themselves were mixed and negative (particularly in the Feyten and Flaitz studies), and it is reassuring that this is reported openly, without exaggeration of findings; only two reported positive results. The overall result, however, is that it is difficult to claim on the basis of these studies that global proficiency can be clearly improved by strategy training. These two studies were of great interest as the numbers of participants were large, in both Spanish and French language learning, and across school and tertiary education levels. They looked at usual end-of-semester test results as measures of the intervention effect, and while their efforts in remaining as naturalistic as possible are clearly worthy, the diversity of settings and participants may have hidden or weakened the detection of what was happening.

It is problematic to combine or synthesise the five studies in a straightforward manner as there are differences between the interventions and settings. However, they all looked at global performance as an outcome, and in this respect a degree of comparability is warranted.

Authors (date)	Strategy training type	Results as reported
Burgos-Kohler (1991)	Mnemonic strategies for vocabulary and overall language ability	Positive effect on vocabulary acquisition and end-of-semester grade scores
Cadierno-Lopez (1992)	Focusing on grammar input rather than language output for better comprehension and language production	Positive effect on comprehension and production
Feyten <i>et al.</i> (1999)	Awareness-raising on strategies (cognitive and metacognitive) on proficiency	Mixed findings (on end-of-semester grades)
Flaitz and Feyten (1996)	Awareness-raising	Mixed findings
Meskill (1991)	Cognitive strategy training/awareness-raising to improve language ability	No significant improvement

4.4.5 Vocabulary ability

The two studies found that training learners to use keywords to help remember and recall vocabulary items were successful, one in Italian and one in Spanish,

the former at school level (girls only) and the latter at university level. However, there are aspects of the nature of the intervention (unclear difference between the two strategies; keyword and elaborated keyword) and the ways in which the studies were carried out (practice effect from testing; extra work done by intervention group) that limit the reliability of the results.

An earlier study by Lawson and Hogben (1998) could not be included as it is unclear whether or not the report refers to the same study and data. (The authors have not responded to efforts to clarify this.) If it is a different study, it might add to the evidence available for the effect of strategy training on vocabulary acquisition.

Perhaps of greater import is the consideration of the long-term effect of the strategy on learner: that is, firstly, whether there is any long-term benefit to what is a labour-intensive way to learn new words (if it is deployed systematically for all vocabulary) and, secondly, whether it prevents learners developing other more effective ways of learning vocabulary (inferencing, generative examples, contextualisation).

The evidence overall is considered to be weak.

Authors (date)	Strategy training type	Results as reported
Lawson and Hogben (1998)	Mnemonic (keyword) and elaborated mnemonic approaches for vocabulary retention and recall	Positive effect of both found on vocabulary outcomes
Burgos-Kohler (1991)	Mnemonic strategies for vocabulary and overall language ability	Positive effect on vocabulary acquisition and end-of-semester grade scores

4.4.6 Listening ability

The overall evidence in favour of strategy training to improve listening comprehension is weak due to methodological characteristics of the studies. It is a pity that the research evidence does not clearly support the instinctive and professional expectations that certain strategies might improve listening ability; predicting, focused listening, understanding redundancy seem intuitively attractive. This said, the studies generally found that the interventions had a positive effect on the learners' listening ability. It has not been shown in the studies that this effect lasts, nor that it was solely due to the intervention. In the case of the Ozeki study in particular, the control group improved considerably and the degree of improvement was more than half that of the intervention groups.

All the studies were in the tertiary education sector and covered the learning of English, Japanese and Russian.

Factors that affected the reliability of the studies include the following:

- only two randomised studies, one of which was unclear randomisation
- selection of particular classes for study
- selection bias in sampling
- reporting inadequacies (and some claims too strong)
- small numbers
- effect of repeated testing

Authors (date)	Strategy training type	Results as reported
McGruddy (1995)	Selective attention for listening comprehension	May be useful
O'Malley <i>et al.</i> (1985)	Mixed package of strategies for listening (and speaking and vocabulary)	Non-significant improvement in listening comprehension
Ozeki (2000)	Integrated strategies for listening comprehension	Considerable improvements in both the control and experimental groups, although greater improvements in the intervention group
Paulauskas (1994)	Prediction, summarising, questioning, and clarifying (plus reciprocal method) for better listening comprehension	No difference between strategies, and strategies plus reciprocal method, but both better than control group
Seo (2000)	Identifying key terms, inferencing, elaborating to improve strategic listening	Positive results in two of the eight tests
Thompson and Rubin (1996)	Mixed strategies (planning, goal defining, monitoring, evaluating, predicting content, listening for redundancy, querying)	Positive results

5. FINDINGS AND IMPLICATIONS

This chapter summarises the findings of the map and synthesis conducted in this review and assesses the implications of these for policy, practice and research.

5.1 Summary of principal findings

This review set out to evaluate the research evidence surrounding the training of language learners to use learning strategies. In terms of effectiveness, clear evidence regarding improvement in overall learning ability would be of great value, particularly if it could be shown to last over an extended period of time. The review focused less on the effectiveness of the various strategies than on the effectiveness of strategy training in general. The relationship between methodological approaches and strength of findings in addressing the review question is also of interest.

Section 3.2.8 reported that 24 of the 38 mapped studies (63%) were of English as a second language (ESL) or English as a foreign language (EFL), and as reported in Chapter 4 on the individual studies at the in-depth evaluation stage, there were nine ESL and six EFL studies (62.5%). It may be argued that these should be examined separately from other languages as there is sufficient difference between learning a language at school or university in the country where that language is spoken as opposed to studying it in a country where it is not used. For this review, the studies were not regarded separately, but it is accepted that in a finer grained examination, and certainly if meta-analysis were feasible, the sub-grouping of analyses (English versus other languages, and even ESL versus EFL) would possibly be warranted. This is less relevant for studies of Spanish being taught in the USA, although we accept the points made implicitly by readers of the draft review concerning English and Spanish that participants of such studies would have access to more Spanish than in countries without such Hispanic presence. There may also be influences on motivation in the cases of studies referred to here, but this has not been explored in any of the studies and no comment can therefore be offered at this stage other than to flag it as a potential issue for consideration. All the studies of Spanish included in the review (N=4) were conducted in the USA and, of the 15 studies of English, five were in the USA (plus two in Canada) while eight were in non-Anglophone countries.

A large number of interventions was found and, as would be expected with the research being carried out around the world at different times by different people, there has been little standardisation of either the packages of interventions or the outcomes that were measured. Some of the research was characterised by the *type* of strategies involved (metacognitive, cognitive and affective) while some focused on the strategies themselves. Of those where the researchers chose to focus explicitly on the type of strategy, metacognitive training appeared to be the most frequent, although, in reality, when categorising the strategies for the review, most by far were strategies of the cognitive type. This may reflect changes of emphasis with no underlying pattern or cause other than differing theoretical conceptualisations and provenance. A degree of standardisation in research method is observed in as much as some of the earliest studies found were a controlled trial and a randomised controlled trial. However, there remains a wide

range of difference in how the controlled studies are constructed (for example, in outcomes of interest and means of evaluation) and reported, and it cannot be suggested that there has been any concerted attempt to unify a research approach.

A consequence of the diversity of approach for the review is that it is no simple matter to combine the findings and doing so is necessarily relatively blunt, with the concomitant risk of trying to compare apples with oranges. Bearing this in mind, 23 of the 25 included studies are considered to be of medium or high relevance in addressing the review question, and it is reassuring to note that at the very least that the research carried out in this field is applied and of interest to the 'real world' of language teaching and learning, rather than of academic interest only. Of course, this does not reflect in any way on the studies which were of lesser relevance to this review as it was not the researchers of the studies who chose the review question!

In terms of reliability, again with the caveat of this being a blunt indicator, 13 of the 25 studies are considered to be of medium reliability, and two of high reliability in addressing the review question. All 25 studies are considered relatively reliable inasmuch as they provide comparative experimental rather than anecdotal evidence, although this does leave 10 studies (40%) that the review found to be of low reliability due to their methodological characteristics.

In considering the overall weight of evidence that the available research provides, factoring together the relevance and reliability indicators, 17 of the 25 studies were considered to be of medium or high weight. In broad terms, this is consistent with a view that there is a considerable amount of solid research evidence to support claims that strategy training for language learning is effective.

However, this must not be over-interpreted, and these indicators say little about the process of the actual interventions and the way in which a particular learning context impinges on the effectiveness or otherwise of an intervention.

It is reassuring that the empirical research provides evidence broadly in line with the theoretical research (which may not surprise some!) and anecdotal evidence. However, it is still vital for users of the research evidence to consider it in the context of their own situations and to weigh up the similarities and differences that should take into account such features as the following:

- level of education/stage of learning
- language in question
- age of learners
- prior learning experience
- generalisability of research findings from one context to another

as well as other considerations that are not covered by the research, such as cost-effectiveness, opportunity cost, and resource availability. Consequently, the compounding of these studies has been limited to relatively blunt amalgamations of findings. However, keeping a perspective on matters, one should note that all the studies included in the review have compared an intervention with something else and have observed, recorded and interpreted the results. In this respect, the studies differ considerably from opinion pieces or theoretical statements on the potential for strategy training.

The variety and composition of interventions, whether single or packaged together in some way, limited the degree to which studies could be combined cumulatively in this review. For example, the two studies on writing, while similar in that they investigated the effect of revision and redrafting of written work, differed in that one was a short intervention over two weeks where the learners read a strategy instruction guide and then redrafted their work using a dictionary, while the other study looked at the effect of redrafting essays over a school year. Both showed strong improvement, but the differences need to be borne in mind and any simple statements on the compounding of these findings would be imprudent.

At the same time, reading strategy interventions allowed more confident 'lumping' as a number of them looked at structural or semantic mapping in relation to comprehension, and all took comprehension as the prime indicator.

A major finding of this review is that none of the studies carried out any long-term, post-intervention testing or follow-up, and it cannot be said for any of the studies, however strong their results, whether the effect of the intervention lasted a week, a month, a year or a lifetime. In terms of cost-effectiveness (regardless of pedagogic effectiveness), this is an important consideration, particularly regarding questions of policy, and the review finds that there is no evidence to support policy decisions in terms of the likelihood of long-term benefits of strategy training; this is, however, very different from finding evidence of no benefit. This is a pity, as in many studies it would have been possible to incorporate follow-up delayed testing of the intervention. A concomitant question that this leaves unanswered is whether or not the beneficial effect of any training can be maintained, reinforced or enhanced by smaller refresher 'doses' of the intervention, say through scaffolding during regular instruction. Such process evaluations would considerably strengthen the value of interventions without great increase in, for example, cost. Research on such a follow-up basis might be achieved, for example, using sampling techniques with qualitative reports from learners which would enhance the findings of interventions carried out using larger scale quantitative methods. Protocols for individual learners would enable researchers to assess the effect of interventions on individuals and complement the findings with detail in the context of the bigger picture.

One study incorporated into its design a degree of longer-term evaluation in that the study took place over a complete year (redrafting of essays) and found a positive effect. It is interesting that corroborative evidence from learners supported the findings but also revealed that the learners preferred the traditional method of not revising their essays!

The study interventions were assessed over a wide variety of outcomes for the main domains of speaking, listening, reading, writing, and so on. Many of the studies used both externally validated tests (Test of English as a Foreign Language (TOEFL), Neale, etc.) and locally constructed tools, questionnaires, examinations, etc. to evaluate the effect of the interventions. Some studies used regular end-of-term or end-of-year tests as indicators, and the trade-off to be considered here is between the desirability of the naturalistic approach of using regular term/year tests and the precision perhaps afforded by specially constructed tests or less naturalistic instruments. Some studies involved self-reporting by the learners and other proxy measures, as triangulation of findings through assessment of strategy use. This was naturally the case concerning perceptions by learners of their learning, strategy use, progress, response, etc.

Interventions varied in length, with the studies looking at awareness-raising typically being shorter than those testing a programme of activities. As reported, the protocol for this review was changed in light of discussion of this issue; awareness-raising can happen in a moment and there is therefore no real need to specify an absolute minimum length or duration of intervention. This is an interesting matter, as it may predicate the possibility that positive benefits could be achieved simply through awareness raising rather than 'invasive' non-naturalistic intervention programmes.

The ways in which studies were carried out, and even more so the ways in which they were reported, varied considerably. Unlike other areas of education, there are for strategy training a relatively high number of controlled and randomised controlled trials. However, there is often a lack of detail that raises queries concerning sampling, details of interventions, characteristics of participants, and the discussion of confounding features in studies. The latter includes the traditional difficulty of controlling for experimenter bias. In some cases, no information is given about the participants and there is scarce information of baseline characteristics, although some testing of homogeneity between groups is carried out in the randomised trials. Very few of the studies report any details on the randomisation procedure and sample selection, and none gives full details of power calculations for sample size. This is of concern as many of the randomised trials used cluster randomisation, which is an approach that allows whole groups to receive an intervention and be compared with groups that do not, but that also then requires larger samples as the power to detect the effect is reduced. Further difficulties in method and reporting include a minimal concern with blinding of allocation to groups, allocation to experimenter, and blinding of assessment of the intervention. These methodological issues are probably being addressed as research methodology skills become more widespread and collaborative research employs the skills of different experts.

The review has not been able to search the non-English language literature systematically, although a number of databases in languages other than English were searched, and several non-English language reports were retrieved (one awaiting translation, and contact with the author made by email). In all but one case, abstracts have been available for the non-English reports identified.

The mainstream publication and dissemination of research based information on strategy training has been through journal articles and books, and this has led to a degree of publication bias as doctoral theses and masters' dissertations have not generally found their way into the knowledge base. McDonough (1999) does, however, cite two unpublished PhD theses. This review has been able to add the findings from a number of postgraduate research studies which may otherwise not have been incorporated into the body of evidence available.

5.2 Strengths and limitations of this systematic review

This is the first systematic review of research evidence in relation to the question. The review brings to light research evidence that was not previously in the mainstream body of knowledge, and it highlights areas of need in terms of future research, research method and quality.

5.2.1 Searching

The time and resources have not been available to do any extensive handsearching of journals, and this remains a long-term objective. The results will be incorporated into updates of the review.

5.2.2 Lumping

Only fairly crude and blunt synthesis of study evidence has been possible due to diversity of intervention type, diversity of outcomes and diversity of measurement instruments. Further time and resources would be needed to bring together numerical data if meta-analysis were desirable and feasible. Some studies present numerical data (means and standard deviations) potentially suited to meta-analysis, but the comparability of included studies would need careful assessment.

5.2.3 Average answer to average question

This review is unable to capture the detail available in rich descriptive reports of individual cases, and is therefore unable to make any statement on the applicability of a specific strategy training intervention for a particular situation. However, the payoff is perhaps in greater generalisability of findings and the ability of the review to find that overall research evidence supports the effectiveness of strategy training in general.

5.2.4 Studies reported in languages other than English

For update of this review, effort will be put into searching for reports of studies published in the non-English language research literature. Time and resources have not yet permitted the systematic searching for such reports.

5.3 Implications

5.3.1 Policy

This review pulls together available research evidence and comes to the conclusion that strategy training is effective. The evidence for its effectiveness is stronger for adult and higher education learners, but there is no systematic picture for the ability level at which training is likely to be most effective, and it is impossible with the evidence currently available to match training interventions to learner need at particular ability levels. This is a pity as it would be useful to find out, for example, whether redrafting of writing (effective at secondary and adult level for whole essays) is also effective at, say, sentence level for beginners at primary level.

Several other aspects remain unclear:

- Is a short sharp awareness training intervention any less effective than training in specific strategy behaviours? The available research does not reveal whether an awareness training programme (potentially less resource-

dependent) for a given strategy or set of strategies would be any less effective than a full training intervention that incorporates implementation of the strategy. This is particularly relevant at Key Stage 2 in the UK where there is a clear intention to introduce language learning strategies (see Key Stage 2 consultation document).

- Is the effectiveness of training related to the combination of strategies in a bundle or package or to certain discrete strategies? Or, in other words, is cognitive strategy training effective or is training in certain cognitive strategies more effective in certain situations? See, for example, work by Cohen (2003): 'What is often lacking is a fine-tuned description of the given strategies, and what may make such descriptions particularly useful would be having them specific to the particular language tasks that the language learners are called upon to perform' (p 1).
- The evidence for the long-term effect of the benefits of strategy training is virtually non-existent. It seems reasonable to assume that, if a strategy training intervention is demonstrated to be effective, the learner somehow incorporates it into their learning mechanisms and that it is compounded along with their other learning experiences and capacities. However, it remains unclear from a research point of view whether this is or is not the case, and the cost-effectiveness of any intervention will remain unsure without longer-term follow-up studies. Particularly from a policy perspective, one assumes it would be unwise to invest in large scale strategy training interventions if the effect could not be demonstrated; on the other hand, low cost (time and money) strategy-awareness training interventions might well be justified until long-term benefit studies provide more evidence.

5.3.2 Practice

Strategy training can be effective. Awareness-raising training interventions and training in implementing the strategies themselves can be shown to have a beneficial effect for learners, but the long-term benefits are unclear, and this has a bearing on the trade-off between the effect and the effort needed to achieve it in terms of time, resources, training, etc.

Notetaking, and semantic and structure mapping are interventions that improve reading comprehension. Most of the studies led to positive findings, but the findings are not necessarily transferable to all pedagogic situations.

Training learners to revise and redraft written work is worthwhile for improving accuracy and quality of output. This can probably be enhanced by training in dictionary skills between drafts 1 and 2.

Some evidence, albeit not strong, is available for the effectiveness of training in strategies to improve oral production (group discussions, accuracy): for example, metacognitive strategies with verbalisation of planning, predicting, monitoring and evaluating and focusing on discrete linguistic items.

Although the findings were mixed, evidence from studies on a variety of packages of strategies shows that overall language ability can be improved by such training. These include keyword and mnemonic strategies for vocabulary; focusing differentially on input and output of linguistic items and strategy awareness raising.

Listening ability can be enhanced by strategy training, with, for example, training in selective attention and other metacognitive strategies. (See also Macaro *et al.*, 2004.)

Potential users of the evidence in this review should conduct a situational analysis to ensure that their own particular context has the characteristics suited to the intervention. Characteristics that are relevant in appraising the potential for effective strategy training include the following:

- age
- stage of learning
- resources available and required
- assessment of outcomes
- training concomitant with or prior to language instruction

The review presents the available evidence for the effectiveness of some strategy training interventions and packages – there are doubtless many others – and these may be more appropriate to a given situation, but may also require detailed evaluation prior to implementation.

5.3.3 Research

It is encouraging that experimental evidence from randomised controlled trials and controlled trials is available to accompany other research evidence. When viewed together, substantial corroboration of research findings should be available.

Non-experimental research needs to be evaluated systematically and incorporated into the body of evidence. This is particularly important in order to understand the full detail of processes in action during strategy training and learner strategy deployment.

More evidence of strategy training for school learners of modern languages is needed, particularly if large-scale training programmes are to be considered.

Different studies had different strengths and weaknesses, and below is a summary of the characteristics that could usefully be addressed to reinforce the reliability of the methods used.

- Clearer randomisation: The procedures used were very rarely reported, and it was often unclear whether individuals, interventions, or experimenters (teachers) had been randomised. This included the lack of reporting, when referring to participant randomisation, on whether randomisation was, for example, applied to the whole sample or sample minus withdrawals at the beginning for example.
- Larger samples of cluster randomisation to ensure the power of the sample to detect the effect of the intervention. Issues of 'leakage' in studies where cluster randomisation was used were not discussed in the vast majority of cases.
- Concealment of allocation (to intervention, to assessor and in assessment, to participants): Greater blinding of assessment would assist in the control of bias, although it is often difficult to do this and maintain a naturalistic setting. Crossover studies might be considered where all participants receive the intervention but at different times.

- Standardisation of testing and assessment instruments, and standardisation of outcome and intervention frameworks would enhance validity of assessment methods, cut costs of research, enable easier aggregation of findings across studies, and possibly enable research funding to be used more efficiently. Validation of testing and measurement tools – or greater use of naturalistic settings and standard tests (end of term, year tests plus externally validated tests or sections of, for example, TOEFL, International English Language Testing System (IELTS), etc.) – would also contribute to a harmonisation of approach in the research community.
- Improved reporting of studies at the individual level is desirable, and this includes practitioner-researchers not being afraid of negative results from experiments; these are equally as valuable as positive findings. Much reporting of demographic details is minimal, as is reporting of baseline measurements prior to experiments. Often complicated statistical analyses are reported but more basic data are unavailable (such as groups scores, descriptions of basis for intervention, with completers, dropouts, attrition, etc. not explained. Improved reporting of studies should also include more systematic coverage of previous research.
- Naturalistic evaluations are desirable – that is, studies that attempt to test the intervention with as little disturbance as possible in the day-to-day patterns of the learners in question. This includes using regular tests and grades/scores to measure the effects, and perhaps crossover trials so all participants receive the same treatment overall.

More effectiveness research is needed, and in particular, long-term post-intervention testing.

More research into the process of how strategy training works is desirable: is it awareness-raising, or the modelling of behaviours for learners to imitate, or both?

6. REFERENCES

6.1 Studies included in map (N = 38) and synthesis (N = 25)

Studies included in the in-depth review are marked with an asterisk (*).

A. Primary reports of included studies

Anderson V (1998) Using multiple intelligences to improve retention in foreign language vocabulary study. Unpublished MA dissertation: St Xavier University, Chicago.

Aninao J (1993) Training high school ESL students to use language-learning strategies. Unpublished PhD thesis: Stanford University, California.

*Ayaduray J and Jacobs G (1997) Can learner strategy instruction succeed? The case of higher order questions and elaborated responses. *System* **25**: 561–70.

Baily C (1996) Unobtrusive computerized observation of compensation strategies for writing to determine the effectiveness of strategy instruction. In: Oxford R (ed.) *Language learning strategies around the world: Cross-cultural perspectives*. Honolulu: University of Hawaii Press, pages 141-150.

*Bimmel P, van den Bergh H, Oostdam R (2001) Effects of strategy training on reading comprehension in first and foreign language. *European Journal of Psychology of Education* **16**: 509–29.

*Bishop G (2001) Using quality and accuracy ratings to quantify the value added of a dictionary skills training course. *Language Learning Journal* **24**: 62–9.

*Burgos-Kohler N (1991) The effect of a selected group of language learning strategies upon language development. Unpublished PhD thesis: University of Texas at Austin.

*Cadierno-Lopez T (1992) Explicit instruction in grammar: a comparison of input based and output based instruction in second language acquisition. Unpublished PhD thesis: University of Illinois at Urbana-Champaign.

*Carrell P, Pharis B, Liberto J (1989) Metacognitive strategy training for ESL reading. *TESOL Quarterly* **23**: 647–78.

*El-Koumy, A-S (1999) Effects of three semantic mapping strategies on EFL students' reading comprehension. In: Annual Meeting of the Teachers of English to Speakers of Other Languages, Chicago 1996. ERIC document ED435193 [Full document available on ERIC but original source details unavailable]

*Feyten C, Flaitz J, LaRocca M (1999) Consciousness raising and strategy use (PHASE II of study). *Applied Language Learning* **10**: 15–38.

*Flaitz J, Feyten C (1996) A two phase study involving consciousness raising and strategy use for foreign language learners. In: Oxford R (ed.) *Language Learning*

Strategies Around the World: Cross-cultural perspectives. Honolulu: University of Hawaii Press, pages 211-226.

Fraser C (1999) Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition* **21**: 225–41.

Halbach A (1999) Integrating strategy training into an existing syllabus: a model. *Neusprachliche Mitteilungen aus Wissenschaft und Praxis* **52**: 104–11.

*Holunga S (1994) The effect of metacognitive strategy training with verbalization on the oral accuracy of adult second language learners. Unpublished PhD thesis: University of Toronto.

*Kitajima R (1997) Referential strategy training for second language reading comprehension of Japanese texts. *Foreign Language Annals* **30**: 84–97.

*Kusiak M (2001) The effect of metacognitive strategy training on reading comprehension and metacognitive knowledge. *EUROSLA Yearbook* **1**: 225–74.

Lam W, Wong J (2000) The effects of strategy training on developing discussions skills in an ESL classroom. *ELT Journal* **54**: 245–55.

*Lawson M, Hogben D (1998) Learning and recall of foreign-language vocabulary: effects of a keyword strategy for immediate and delayed recall. *Learning and Instruction* **8**: 179–94.

*McGruddy R (1995) The effect of listening comprehension strategy training with advanced level ESL students with students of English as a second language. Unpublished PhD thesis: Georgetown University, Washington DC.

*McGuire K (1999) Generative Précising as a Reading Comprehension Strategy for Adult ESL Learners. Unpublished PhD thesis: University of California, Los Angeles.

*Meskill C (1991) Language learning strategies advice: a study on the effects of on-line messaging. *System* **19**: 277–87.

*Najar R (1997) The effect of notetaking strategy instruction on comprehension in ESL texts. Unpublished PhD thesis: University of Hawaii.

*O'Malley J, Chamot A, Stewner-Manzanares G, Russo R, Kupper L (1985) Learning strategy applications with students of English as a second language. *TESOL Quarterly* **19**: 557–84.

*Ozeki N (2000) Listening strategy instruction for female EFL college students in Japan. Unpublished PhD thesis: University of Pennsylvania, Indiana.

*Paulauskas S (1994) The effects of strategy training on the aural comprehension of L2 adult learners at the high beginning/low intermediate proficiency level. Unpublished PhD thesis: University of Toronto.

*Raymond P (1993) The effects of structure strategy training on the recall of expository prose for university students reading French as a second language. *The Modern Language Journal* **77**: 445–58.

Ridley J, Singleton D (1995) Strategic L2 lexical innovation: case study of a university-level ab initio learner of German. *Second Language Research* **11**: 137–48.

Riley L, Harsch K (1999) Enhancing the learning experience with strategy journals: supporting the diverse learning styles of ESL/EFL students. HERDSA Annual International Conference, Melbourne. Available at: <http://www.herdsa.org.au/branches/vic/Cornerstones/pdf/Riley.PDF>

Salataci R, Akyel A (2002) Possible effects of strategy instruction on L1 and L2 reading. *Reading in a Foreign Language* **14**: 1–17. Available at: <http://nflrc.hawaii.edu/rfl/April2002/salataci/salataci.html>

*Sengupta S (2000) An investigation into the effects of revision strategy instruction on L2 secondary school learners. *System* **28**: 97–113.

*Seo K (2000) Intervening in tertiary students' strategic listening in Japanese as a foreign language. Unpublished PhD thesis: Griffith University, Australia.

Simmons D (1996) A study of strategy use in independent learners. In: Pemberton R (ed.) *Taking Control: Autonomy in Language Learning*. Hong Kong: Hong Kong University Press, pages 61–75.

Song M (1997) Developing reading strategies through reciprocal teaching method. *Language Research* **33**: 321–34.

Stokes D (1981) The effects of various discrete-point focusing strategies on performance in Spanish. Unpublished PhD thesis: Indiana University.

*Talbot D (1995) Metacognitive strategy training for reading: developing second language learners' awareness of expository text patterns. Unpublished PhD thesis: University of Hong Kong.

Tang H, Moore D (1992) Effects of cognitive and metacognitive pre-reading activities on the reading comprehension of ESL learners. *Educational Psychology* **12**: 315–31.

*Thompson I, Rubin J (1996) Can strategy instruction improve listening comprehension? *Foreign Language Annals* **29**: 331–42.

B. Secondary reports of included studies

Najar R (1999) The effect of cognitive learning strategy instruction: an EFL classroom study. *AARE 1999 Conference Report*. Full document available at <http://www.aare.edu.au/99pap/naj99035.htm> [Clarification has been requested by email from the author. In the absence of a response, this report and Najar (1997) have been assumed to relate to one study only.]

Hogben, D & Lawson, M. J (1994) Keyword and multiple elaboration strategies for vocabulary acquisition in foreign language learning. *Contemporary Educational Psychology* **1**: 204–13. [Clarification has been requested by email from the authors. In the absence of a response, this report and Lawson and Hogben (1998) have been assumed to relate to one study only.]

6.2 Studies excluded from in-depth review, with reasons for exclusion

A. No comparison included in the study

Anderson V (1998) Using multiple intelligences to improve retention in foreign language vocabulary study. Unpublished MA dissertation: St Xavier University, Chicago. [action research].

Aninao J (1993) Training high school ESL students to use language-learning strategies. Unpublished PhD thesis: Stanford University, California. [descriptive].

Fraser C (1999) Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition* **21**: 225–41. [interrupted time series].

Halbach A (1999) Integrating strategy training into an existing syllabus: a model. *Neusprachliche Mitteilungen aus Wissenschaft und Praxis* **52**: 104–11. [descriptive].

Lam W, Wong J (2000) The effects of strategy training on developing discussions skills in an ESL classroom. *ELT Journal* **54**: 245–55. [descriptive].

Ridley J, Singleton D (1995) Strategic L2 lexical innovation: case study of a university-level ab initio learner of German. *Second Language Research* **11**: 137–48. [case study].

Salataci R, Akyel A (2002) Possible effects of strategy instruction on L1 and L2 reading. *Reading in a Foreign Language* **14**: 1–17. Available at: <http://nflrc.hawaii.edu/rfl/April2002/salataci/salataci.html> [descriptive].

Simmons D (1996) A study of strategy use in independent learners. In: Pemberton R (ed.) *Taking Control: Autonomy in Language Learning*. Hong Kong: Hong Kong University Press. [ethnographic approach].

Song M (1997) Developing reading strategies through reciprocal teaching method. *Language Research* **33**: 321–34. [descriptive].

Stokes D (1981) The effects of various discrete-point focusing strategies on performance in Spanish. Unpublished PhD thesis: Indiana University. [trial].

Tang H, Moore D (1992) Effects of cognitive and metacognitive pre-reading activities on the reading comprehension of ESL learners. *Educational Psychology* **12**: 315–31. [descriptive].

B. Studies which do not test the effect of the intervention

Baily C (1996) Unobtrusive computerized observation of compensation strategies for writing to determine the effectiveness of strategy instruction [controlled trial].

Riley L, Harsch K (1999) Enhancing the learning experience with strategy journals: supporting the diverse learning styles of ESL/EFL students [descriptive, comparative study].

6.3 Studies identified but retrieved after the cut-off (N = 14)

The study marked with an asterisk (*) is awaiting translation.

Avila E, Sadoski M (1996) Exploring New Applications of the Keyword Method to Acquire English Vocabulary. *Language Learning* **46**: 379–95.

Aziz L (1995) A Model of Paired cognitive and metacognitive strategies: its effect on second language grammar and writing performance. Unpublished PhD thesis: University of California, Los Angeles.

*Bimmel P, Oostdam R (1998) Strategietraining en leesvaardigheid. *Levende Talen [Strategy training and reading skills]* **543**: 556–64.

Cohen A, Weaver S, Li T-Y (1998) The impact of strategies-based instruction on speaking a foreign language. In: Cohen AD (ed.) *Strategies in learning and using a second language*. Harlow: Longman, pages 107–56.

Dadour S, Robbins J (1996) University-level studies using strategy instruction to improve speaking ability in Egypt and Japan. In: Oxford R (ed.) *Language Learning Strategies Around the World: Cross-cultural perspectives*. Honolulu: University of Hawaii Press, pages 157-166.

Klohs L (1994) Use of mnemonic strategies to facilitate written production of a second language by high school French students. Unpublished PhD thesis: University of Minnesota.

Lin T-C (1995) Effects of prompted self-elaborations with embedded strategic cues on second language learners in a hypermedia environment. Unpublished PhD thesis: Purdue University, Indiana.

Macaro E (2001) Learning strategies in foreign and second language classrooms: an intervention study in detail. In: Macaro E (ed.) *Learning Strategies in Foreign and Second Language Classrooms*. London: Continuum, pages 128–74.

Macaro E, Vanderplank R, Graham S, Nakatani Y, Wingate W (2004) A systematic review of effective teaching approaches for uni-directional listening comprehension for learners of a modern language. Protocol for a systematic review. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. Available at: http://eppi.ioe.ac.uk/EPPIWebContent/reel/review_groups/MFL/

Nakatani Y (2002) Improving Oral Proficiency Through Strategy Training. Unpublished PhD thesis: University of Birmingham.

Nunan D (1997) Strategy training in the language classroom: an empirical investigation. *RELC Journal* **28**: 56–81.

Oxford R (1994) Language learning strategies. In: *Language Learning Strategies: An update*. ERIC Digest. Washington, DC: Office of Educational Research and Improvement.

Oxford R, Crookall D, Cohen A, Lavine R, Nyikos M, Sutter W (1990) Strategy training for language learners: six situational case studies and a training model. *Foreign Language Annals* **22**: 197–216.

Rubin J, Thompson I (1992) Materials Selection in Strategy Instruction for Russian Listening Comprehension. ERIC document ED349796. [Full document available on ERIC but original source details unavailable]

Wardrop E, Anderson K (1992) An investigation into learners' reactions to learner training for self-access listening, *Edinburgh Working Papers in Linguistics* **3**: 135–45.

6.4 Studies identified after this report's first draft

Four additional potential studies that may meet the map or in-depth review inclusion criteria have come to light. These have been received beyond the cut-off date for consideration in this version but will be incorporated into its first update where appropriate. We are in any case grateful to the authors of these studies for making us aware of their work.

Erler L, Macaro E (2002–2004) Learning to learn: strategy training in reading and writing as a foreign language. ESRC project (ongoing).

Ikeda M, Takeuchi O (2003) Can Strategy Instruction Help EFL Learners to Improve their Reading Ability? An Empirical Study. *JACET Bulletin* **37**: 49–60.

Paige M, Cohen A, Shively R (forthcoming) Assessing the Impact of a Strategies-Based Curriculum on Language and Culture Learning Abroad. *Frontiers: The interdisciplinary Journal of Study Abroad*.

Takeuchi O, Wakamoto N (2001) Language learning strategies used by Japanese college learners of English: a synthesis of four empirical studies. *Language Education and Technology* **38**: 21–43.

6.5 Other references used in the text of the report

Baird J, White R (1984) *Improving learning through enhanced metacognition: A classroom study*. Paper presented at the 68th meeting of the American Educational Research Association, New Orleans, LA, USA: 23–27 April.

Benson P (2005) *Autonomy, independence and strategies* bibliography. Available at <http://ec.hku.hk/autonomy/bibliog.html>

Benson P (1995) A critical view of learner training. *Learning Learning* **2:2** (electronic journal). Available at <http://penta.ufgrs.br/edu/telelab/7/phil22e.htm>

Biggs J (1987) *Student Approaches to Learning and Studying*. Melbourne: Australian Council for Educational Research.

Cajkler W, Thornton B (1999) Language learner perceptions of strategy use in secondary schools. *Language Learning Journal* **20**: 45–50.

- Caroll JB (1973) Implications of aptitude test research theory and psycholinguistic theory for FL teaching. *International Journal of Psycholinguistics* 2: 5–14.
- Carton A (1971) Inferencing: a process in using and learning language. In: Pimsleur P, Quinn T (eds) *The Psychology of Second Language Learning*. Cambridge: Cambridge University Press, pages 45-58.
- Chamot A (1987) The learning strategies of ESL students. In: Wenden A, Rubin J (eds) *Learner Strategies in Language Learning*. Eaglewood Cliffs, New Jersey: Prentice-Hall International, pages 71–84.
- Chamot A, O'Malley M (1996) The cognitive academic language learning approach: a model for linguistically diverse classrooms. *Elementary School Journal* 96: 259–73.
- Cohen A (1998) *Strategies in Learning and Using a Second Language*. London and New York: Longman.
- Cohen A (2003) Strategy-based learning of speech acts: developing and evaluating a web-based curriculum. Paper presented at the Independent Language Learning Conference. The Open University, Milton Keynes, England: 4–5 December.
- Correa M (1995) Incorporating Cooperative Learning Strategies To Improve Science Achievement Scores among Ninth Grade ESOL I and II Physical Science Students. EdD Practicum Report: Nova Southeastern University, Florida.
- DfES (2002) *The Green Paper 14–19: extending opportunities, raising standards*. Nottingham: DfES Publications. Available at: <http://www.dfes.gov.uk/14-19greenpaper>
- DfES (2003a) *Framework for Teaching Modern Foreign Languages: Years 7, 8 and 9*. Nottingham: DfES Publications. Available at: http://www.standards.dfes.gov.uk/keystage3/strands/publications/?template=down&pub_id=2463&strand=tf
- DfES (2003b) *Languages for All: Languages for Life A Strategy for England*. Nottingham: DfES Publications. Available at: <http://www.dfes.gov.uk/languagesstrategy/pdf/DfESLanguagesStrategy.pdf>
- DfES (2004) Key Stage 2 MFL framework. Unpublished consultation document. London: DfES. Available at: <http://www.dfes.gov.uk/consultations/conResults.cfm?consultationId=1265>
- Dickinson L (1992) *Learner Autonomy 2: Learner training for language learning*. Dublin: Authentik Language Learning Resources.
- Ellis G, Sinclair B (1989) *Learning to Learn English*. Cambridge: Cambridge University Press.
- EI-Koumy A-S (1997). Effects of three questioning strategies on EFL reading comprehension. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages (30th, Chicago, IL, March 1996).
- EPPI-Centre (2002a). *Core Keywording Strategy, Version 0.9.5*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

- EPPI-Centre (2002b) *Review Guidelines for Extracting Data and Quality Assessing Primary Studies in Educational Research, Version 0.9.5*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Eysenck M, Keane M (2001). *Cognitive Psychology*. Hove: Psychology Press.
- Gooden-Jones E (1996) Developing writing proficiency through cooperative learning strategies in limited-English-proficient college students. PhD thesis: Digital Dissertations, Proquest (<http://wwwlib.umi.com/dissertations/>).
- Graham S, Macaro E, Nakatani Y, Vanderplank R, Wingate U (2004) *A systematic review of effective teaching approaches for uni-directional listening comprehension for learners of a modern language*, protocol for an EPPI-Centre systematic review, http://eppi.ioe.ac.uk/EPPIWebContent/reel/review_groups/MFL/ML_protocol2.pdf
- Harris V (1997) *Teaching Learners How to Learn: Strategy training in the ML classroom*. London: CILT.
- Ishii S, Klopff D (1980) Increasing instructional effectiveness: reducing classroom apprehension. Paper presented at the International Conference of the Japan Association of Language Teachers on Language Teaching/Learning. Nagoya, Japan: 22–24 November.
- Jo V (1997) Effects of Instructional Strategies on Second Language Acquisition Processes. *Working Papers in Educational Linguistics* 13: 51–60.
- Kaelin A (1991) The Effects of Instruction Using a Mnemonic Graphic Organizer on vocabulary Acquisition among Adult English-as-a-Second-Language Students. Unpublished Master's thesis: California State University, Sacramento.
- Kern R (1989). Second language reading strategy instruction: its effects on comprehension and word inference ability. *The Modern Language Journal* 73: 135–49.
- Kresovich B (1990) Timed writing practice for Japanese University students in an EFL setting. *Bulletin of College of Education* 37: 121–28.
- Lane T, Potter B (1998) Teaching collaborative feedback strategies in intermediate writing. Paper presented at the 32nd Annual Meeting of the Teachers of English to Speakers of Other Languages. Seattle, WA, USA: 17–21 March.
- Luk V (1992) Interactional listening tasks: a comparative study of strategy and practice teaching approaches. *Edinburgh Working Papers in Applied Linguistics* 3: 78–91.
- Macaro E (1997) *Target Language, Collaborative Learning and Autonomy*. Clevedon: Multilingual Matters Ltd.
- Macaro E (2001) *Learner Strategies in Second and Foreign Language Classrooms*. London: Continuum.

Macaro E, Vanderplank R, Richards B (2003) Strategy training in year 12 French: Listening and Writing. [Ongoing research reported in CERUK: see <http://195.194.2.38/starweb/ceruk/servlet.starweb>]

McDonough S (1999) State of the art article: learner strategies. *Language Teaching* **32**: 1–18.

Miyake A, Shah P (Eds) (1999) *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*. New York: Cambridge University Press.

Moulden (1981) Une experience d'apprentissage autodirigé de l'anglais dans un cadre institutionnel. *Etudes de Linguistique Appliquée* **41**: 24–38.

The Nuffield Languages Inquiry (2000) *Languages: the next generation. Final report and recommendations of the Nuffield Languages Inquiry*. London: The Nuffield Foundation

O'Malley M (1987) The effects of training in the use of learning strategies on acquiring English as a second language. In: Wenden A, Rubin J (eds) *Learner Strategies in Language Learning*. New York: Prentice Hall, pages 133-144.

O'Malley M, Chamot A (1990) *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

Oxford R (1986) *Second Language Learning Strategies: Current research and implications for practice*. Los Angeles: California University, Center for Language Education and Research

Oxford R (1990) *Language Learning Strategies: What Every Teacher Should Know*. Rowley, Massachusetts: Newbury House.

Oxford R, Crookall D (1989) Research on language learning strategies: methods, findings, and instructional issues. *Modern Language Journal* **73**: 4–19.

Pachler N (2002) Foreign language learning in England in the 21st Century. *Language Learning Journal* **25**: 4–7.

Pachler N (2003) Foreign language teaching as an evidence-based profession? *Language Learning Journal* **27**: 4–14.

Perez C (1990) How the knowledge of rhetorical structure helps improve reading and writing skills. Paper presented at the World Congress of Applied Linguistics sponsored by the International Association of Applied Linguistics. Thessaloniki, Greece: 15–21 April.

Politzer R, Weiss L (1969) *Experiment in Improving Achievement in FL through Learning Selected Skills*. Stanford, California: Stanford University.

Rees-Miller J (1993) A critical appraisal of learner training: theoretical bases and teaching implications. *TESOL Quarterly* **27**: 679–89.

Rubin J (1975) What the good language learner can teach us. *TESOL Quarterly* **9**: 41–51.

- Rubin J (1981) The study of cognitive processes in second language learning. *Applied Linguistics* **2**: 117–31.
- Sinclair B (1999) More than an act of faith? Evaluating learner autonomy. In: Kennedy C (ed.) *Innovation and Best Practice*. Harlow: Longman; in association with the British Council, pages 96–107.
- Singhal M (1998) Improving reading comprehension through strategy instruction. *Arizona Reading Journal* **25**: 31–39.
- Smith P (1985) It's not their fault. *Journal of Language Studies* **2**: 42–59.
- Starks-Martin G (1996) Using "Think-Alouds," journals, and portfolios to assess among students' perceptions of their study/learning strategies. Paper presented at the 30th Annual Meeting of the teachers of English to Speakers of Other Languages. Chicago, IL, USA: 26–30 March.
- Stern H (1975) What can we learn from the good language learner? *Canadian Modern Language Review* **31**: 304–18.
- Viswat L, Jackson S (1994) The effect of strategies training on student errors on a listening cloze. *Journal of the Himeji Dokkyo University College of Foreign Languages*: 235–49.
- Weatherford HJ (1990) Techniques for learning vocabulary. Paper presented at the Annual Meeting of the Foreign Language Association of Georgia. Savannah, GA, USA: 10 February.
- Weinstein C, Mayer R (1986) The teaching of learning strategies. In: Wittrock MC (ed.) *Handbook of Research on Teaching, 3rd Edition*. New York: Macmillan, pages 315–27.
- Wenden A, Rubin J (eds) (1987) *Learner Strategies in Language Learning*. New York: Prentice Hall.
- Yang N-D (1995) Effective awareness-raising in language learning strategy training. Paper presented at the 29th Annual Meeting of the Teachers of English to Speakers of Other Languages. Long Beach, CA, USA: 26 March – 15 April.
- Yang N-D (1998) Exploring a role for teachers: promoting learner autonomy. *System* **26**: 27–35.
- Zhicheng Z (1992) The effects of teaching reading strategies on improving reading comprehension for ESL learners. Paper presented at the Annual Meeting of the Mid-South Educational Research Association. Knoxville, TN, USA: 11–13 November.

Appendix 1.1: Advisory Group membership

Assistance with reviewing

Concha Furnborough The Open University, UK

Yasuo Nakatani Oxford University

Lynn Erler Oxford University

Searching, document retrieval and administration

Sue Walters, Oxford University

External advisor

Lynn Erler Oxford University

Appendix 2.1: Inclusion and exclusion criteria

To be included in the map, reports needed to include *all* the following:

1. a strategy training intervention in language learning
2. an intervention carried out in a formal setting such as groups of learners in schools, universities and language centres
3. a study not primarily involving bilingual learners
4. primary, empirical research
5. research carried out since 1960

Appendix 2.2: Search strategy for electronic databases

Databases searched

BEI (British Educational Index)
CERUK (Current Educational Research in the UK)
Dissertation Abstracts
ERIC
REEL
SPECTR (Social, Psychological and Educational Controlled Trials Register)
PsycINFO
Linguistics and Language Behavior Abstracts
Educational Administration Abstracts
ISI Citation Indexes
Mental Measurements Yearbook
MLA International Bibliography
UNESBIB (UNESCO Bibliographic Database)
UNESDOC (UNESCO documents collection)
IAED (International Archive of Education Data)
PAIS (Public Affairs Information Service)
CILT (Centre for Information on Language Teaching)

Terms used in database searches

The following terms for strategies, strategy learning, or strategy training were used in database searches (using database controlled terms where possible):

1. affective-strateg*
2. autonomie-guidée
3. autonomisation
4. autonomous-learning
5. cognitive-strateg*
6. cognitive-style
7. language-learning-strateg*
8. learner-autonomy
9. learner-strateg*
10. learner-train*
11. learner-based-teaching
12. learn*-style*
13. lernerautonomie
14. meta-cognit*-strateg*
15. self-directed-learn*
16. self-managed-learning
17. self-instruction
18. strateg*
19. strategies-based-instruction
20. strateg*-training
21. student-autonomy
22. student-centred-learning

Appendix 2.4: EPPI-Centre keyword sheet, including review-specific keywords

V0.9.7 Bibliographic details and/or unique identifier

<p>A1. Identification of report Citation Contact Handsearch Unknown Electronic database (Please specify.)</p> <p>A2. Status Published In press Unpublished</p> <p>A3. Linked reports <i>Is this report linked to one or more other reports in such a way that they also report the same study?</i></p> <p>Not linked Linked (Please provide bibliographical details and/or unique identifier.) </p> <p>A4. Language (Please specify.) </p> <p>A5. In which country/countries was the study carried out? (Please specify.) </p>	<p>A6. What is/are the topic focus/foci of the study? Assessment Classroom management Curriculum* Equal opportunities Methodology Organisation and management Policy Teacher careers Teaching and learning Other (Please specify).....</p> <p>A7. Curriculum Art Business studies Citizenship Cross-curricular Design and technology Environment General Geography Hidden History ICT Literacy – first language Literacy further languages Literature Maths Music PSE Physical education Religious education Science Vocational Other (Please specify.).....</p>	<p>A8. Programme name (Please specify.) </p> <p>A9. What is/are the population focus/foci of the study? Learners Senior management Teaching staff Non-teaching staff Other education practitioners Government Local education authority officers Parents Governors Other (Please specify.).....</p> <p>A10. Age of learners (years) 0–4 5–10 11–16 17–20 21 and over</p> <p>A11. Sex of learners Female only Male only Mixed sex</p>	<p>A12. What is/are the educational setting(s) of the study? Community centre Correctional institution Government department Higher education institution Home Independent school Local education authority Nursery school Post-compulsory education institution Primary school Pupil referral unit Residential school Secondary school Special needs school Workplace Other educational setting (Please specify.).....</p> <p>A13. Which type(s) of study does this report describe? A. Description B. Exploration of relationships C. Evaluation a. naturally-occurring b. researcher-manipulated D. Development of methodology E. Review a. Systematic review b. Other review</p>
--	--	--	---

Review-specific questions

The following questions were used to augment the generic EPPI-Centre data extraction questions.

Additional questions about the strategy training that was evaluated

A.1 Have the strategies been defined (if X and Y then Z) by the researcher/teacher?

A.2 Is the way the strategy is supposed to lead to learning or improved proficiency clear?

A.3 Is there a clear relationship between the strategy training (what the teacher did with the students) and what the students would be expected to do eventually as independent individual learners?

A.4 Has the intervention been made explicit to the reader (even if it is not necessarily explicit to the learners)? Consider the way it was carried out; length of time; number of repeated exposure to a strategy; whether scaffolded or unscaffolded; monitored or unmonitored; evaluated.

A.5 Is there an investigation as to change in strategic behaviour as a result of the intervention?

Additional questions about outcome measurement

A.6 Is there an attempt to collect data as to (at least) associative evidence between strategy training and proficiency gains? Or, if not, is there an attempt to collect data as to (at least) associative evidence between strategy training and another outcome?

A.7 Have delayed post-tests been carried out after a period of withdrawal from strategy training?

A.8 Do the post-tests measure both strategic behaviour and other variables, such as proficiency gains or motivation?

A.9 Is it clear that there is no equivalence between the strategies used in the training and the testing method used in the proficiency measures in order not to favour the experimental group?

A.10 Is there any triangulation (other than in A.5 above) on the effectiveness of strategy training on proficiency (for example, by asking the learners' opinions)?

A.11 Does the strategy training time come out of normal teaching time?

Appendix 3.1: Details of studies included in the systematic map

Studies by outcomes measured

Outcome	Name of author(s) and year of publication
Asking and answering higher order questions	Ayaduray and Jacobs (1997)
Attitude	Sengupta (2000)
Aural ability	McGruddy (1995), O'Malley <i>et al.</i> (1985)
Awareness	Holunga (1994), Kusiak (2001), Talbot (1995)
Comprehension	Bimmel <i>et al.</i> (2001), Cadierno-Lopez (1992), Carrel <i>et al.</i> (1989), El-Koumy (1999), Kitajima (1997), Kusiak (2001), McGruddy (1995), McGuire (1999), Najar (1997), Ozeki (2000), Paulauskas (1994), Salataci and Akyel (2002), Seo (2000), Song (1997), Talbot (1995), Tang and Moore (1992), Thompson and Rubin (1996)
Global proficiency	Burgos-Kohler (1991), Feyten <i>et al.</i> (1999), Flaitz and Feyten (1996), Kusiak (2001), Meskill (1991), Seo (2000), Stokes (1981)
Grammar accuracy	Aninao (1993), Cadierno-Lopez (1992), Holunga (1994), Kitajima (1997), McGuire (1999)
Interpreting meaning	Cadierno-Lopez (1992)
Metacognitive knowledge	Kusiak (2001)
Not stated	Halbach (1999)
Oral ability	Lam and Wong (2000), O'Malley <i>et al.</i> (1985)
Recall (content/meaning units)	Lawson and Hogben (1998), Raymond (1993)
Self-perception of ability	Kusiak (2001), Meskill (1991)
Strategy awareness	Holunga (1994)
Strategy transfer	Bimmel <i>et al.</i> (2001)
Strategy use	Aninao (1993), Baily (1996), Flaitz and Feyten (1996), Holunga (1994), McGruddy (1995), McGuire (1999), Ozeki (2000), Riley and Harsch (1999), Simmons (1996)
Vocabulary acquisition	Anderson (1998), Aninao (1993), Burgos-Kohler (1991), Fraser (1999), Lawson and Hogben (1998), Ridley and Singleton (1995)
Writing ability (accuracy and quality)	Bishop (2001), Sengupta (2000)

Studies by education sector

Education sector	Name of author(s) and year of publication
Adult	Baily (1996), Holunga (1994), McGuire (1999), Meskill (1991), Paulauskas (1994), Simmons (1996)
Higher education (non-university)	Ozeki (2000), Riley and Harsch (1999), Talbot (1995)
Secondary school	Anderson (1998), Aninao (1993), Ayaduray and Jacobs (1997), Bimmel <i>et al.</i> (2001), Feyten <i>et al.</i> (1999), Kusiak (2001), Lam and Wong (2000), Lawson and Hogben (1998), O'Malley <i>et al.</i> (1985), Sengupta (2000), Tang and Moore (1992)
University	Bishop (2001), Burgos-Kohler (1991), Cadierno-Lopez (1992), Carrel <i>et al.</i> (1989), El-Koumy (1999), Feyten <i>et al.</i> (1999), Flaitz and Feyten (1996), Fraser (1999), Holunga (1994), Kitajima (1997), McGruddy (1995), Najar (1997), Raymond (1993), Ridley and Singleton (1995), Salataci and Akyel (2002), Seo (2000), Song (1997), Stokes (1981), Tang and Moore (1992), Thompson and Rubin (1996)

Appendix 4.1: Details of studies included in the in-depth review (N = 25)

Author, year, method, publication type	Study title	Setting, population (N = intervention/ total)	Strategy training intervention	Domain, intervention duration	Outcomes measured	Strategy change measured?	Delayed testing?	Results as reported
Ayaduray and Jacobs, 1997 RCT (cluster) Article	Can learner strategy instruction succeed? The case of higher order questions and elaborated responses	Singapore ESL Secondary school N = 16/32 in two groups	Instruction in asking higher order questions	Speaking 10 wks	(1) Frequency of asking higher order Qs (2) Elaborated responses	No	No	<i>Positive</i> Learners become better questioners and group discussion improves.
Bimmel <i>et al.</i> , 2001 Comparative study Article	Effects of strategy training on reading comprehension in first and foreign language	The Netherlands Secondary school EFL (and Dutch) N = 12/119 in matched pairs	Looking for key fragments Paying attention to structure Making up questions Mapping most important information	Reading 15 weeks	(1) Use of strategies in Dutch and comprehension in Dutch (2) Reading comprehension in English (strategy training transfer from first to second language)	No	No	<i>Negative</i> Strategy training works for first language but does not transfer to second.
Bishop, 2001 Comparative study Article	Using quality and accuracy ratings to quantify the value added of a dictionary skills training course	UK French Adult University N = 15/30 in two groups	Using dictionary between essay draft and redraft	Writing 10 hours	(1) Essay length (2) Vocabulary usage (3) Vocabulary acquisition	No?	No	<i>Positive</i> >10% improvement in writing

Author, year, method, publication type	Study title	Setting, population (N = intervention/ total)	Strategy training intervention	Domain, intervention duration	Outcomes measured	Strategy change measured?	Delayed testing?	Results as reported
Burgos-Kohler, 1991 RCT (cluster) PhD	The effect of a selected group of language learning strategies upon language development (foreign language instruction)	USA Spanish University N = 104/143 in six groups (two experimental + two control)	Language learning strategies (keyword, elaboration, association, grouping, recombination, contextualisation)	Vocabulary achievement, Overall performance Six weeks x 10 minutes	(1) Semester grades (2) Vocabulary usage (3) Vocabulary acquisition	No	No	<i>Positive</i> Improvement in vocabulary and improvement in overall language
Cadierno-Lopez, 1992 RCT (cluster) PhD	Explicit instruction in grammar: a comparison of input-based and output-based instruction in second language acquisition	USA Spanish University N = 94/141 in four experimental + two control groups	Input versus output based instruction	Proficiency (comprehension, production, grammar)	(1) Interpreting meaning (2) Producing correct forms	No	One week, one month	<i>Mixed</i> Input- and output-based instruction showed better results than no instruction, but no difference between two interventions.
Carrell <i>et al.</i> , 1989 Comparative study Article	Metacognitive strategy training for ESL reading	USA English University N = 18/26 in four groups (two of each)	Semantic mapping of texts versus experience-text relationship versus nothing	Reading Four days	Reading comprehension	No	No	<i>Mixed</i> Both types of mapping led to better reading comprehension, but no overall difference between them.
El-Koumy, 1999 RCT Report	Effects of three semantic mapping strategies on EFL students' reading comprehension	Egypt English University N = 60/187 in three intervention groups	Student mediated semantic mapping	Reading Five months (20 x one hour)	Reading comprehension	No	No	<i>Positive</i> Teacher-student mediated mapping produced better scores on TOEFL reading comprehension measures.

Author, year, method, publication type	Study title	Setting, population (N = intervention/ total)	Strategy training intervention	Domain, intervention duration	Outcomes measured	Strategy change measured?	Delayed testing?	Results as reported
Feyten <i>et al.</i> , 1999 RCT (cluster) Article	Consciousness raising and strategy use	USA French and Spanish University, middle and high school N = 863	Handout on 26 learning strategies (MAR) or handout on 26 reasons for studying a language (CAR)	Awareness – metacognitive and cognitive strategies 50 mins	Term exam grades	No	No	<i>Mixed</i> Greater positive effect in control group at university level Greater positive effect at high school level MAR and CAR significantly positive effect at middle school level
Flaitz and Feyten, 1996 RCT (cluster) Report/Chapter	A two-phase study – phase I	USA University Spanish N = 130/229 in six experimental and six control groups	Metacognitive strategy awareness	Proficiency 50 mins	(1) Semester grades (2) Strategy use	No	No	<i>Positive</i> A short, sharp dose of awareness-training produces better results on term's exams.
Holunga, 1994 RCT (pairs) PhD	The effect of metacognitive strategy training with verbalisation on the oral accuracy of adult second language learners	Canada ESL Adult N = 32/48 in two intervention and one control group of pairs matched for gender	Metacognitive strategies (predicting, planning, monitoring, evaluation) with and without verbalisation	Speaking Three weeks (15 hours)	(1) Verb use accuracy (2) Awareness of strategies	No	Yes	<i>Positive</i> Verb use accuracy improved, but not at delayed testing.

Author, year, method, publication type	Study title	Setting, population (N = intervention/ total)	Strategy training intervention	Domain, intervention duration	Outcomes measured	Strategy change measured?	Delayed testing?	Results as reported
Kitajima, 1997 Comparative study Article	Referential strategy training for second language reading comprehension of Japanese texts	USA University Japanese N = 28/43 in two groups	Recognising syntactic and discursal links in text	Reading 15 wks/4 per week	(1) Identifying referents (2) Reading comprehension	No	No	<i>Mixed</i> On 2, 3 and 4 of the paragraphs used for testing, the intervention produced better comprehension of texts; on paragraphs 1 and 2, better recognition of referential links.
Kusiak, 2001 Comparative study Article	The effect of metacognitive strategy training on reading comprehension and metacognitive knowledge	Poland English Secondary school 78/158 in two groups	Awareness in reading strategies – finding main idea, recognising topic sentences, text patterns, keywords, guessing meaning. Plus observing others and own learning	Awareness Four weeks (8 x 45 minutes)	(1) Metacognitive knowledge (2) Self evaluation of reading skills (3) Reading comprehension test (4) General language competence	No	No	<i>Positive</i> Metacognitive training is effective for reading comprehension on intermediate learners.
Lawson and Hogben, 1998 Comparative study Article	Learning and recall of foreign language vocabulary: effects of keyword strategy for immediate and delayed recall	Australia High school Female students Italian	Keyword (elaborated) method for vocabulary recall	Vocabulary 10 days/ 3 tests 2 days x 45 mins	Vocabulary recall	No	Yes	<i>Positive</i> Statistically significant positive effect of intervention on vocabulary recall

Author, year, method, publication type	Study title	Setting, population (N = intervention/ total)	Strategy training intervention	Domain, intervention duration	Outcomes measured	Strategy change measured?	Delayed testing?	Results as reported
McGruddy, 1995 Comparative study PhD	The effect of listening comprehension strategy training with advanced level ESL students	USA ESL University Advanced 10/32 in one intervention and two comparison groups	Predicting, inferring and selective attention for listening comprehension	Listening 14 wks at 100 mins/wk	(1) Aural proficiency (2) Comprehension (3) Self-reported strategy use	Yes	No	<i>Positive</i> Significant differences reported, students reported perceived benefit of the training
McGuire, 1999 RCT (cluster) PhD	Generative precising as a reading comprehension strategy for adult ESL learners	USA ESL Adult (private) ESL54/71 in three intervention and one control groups	Generative précising (writing notes next to text...) versus two other 'strategies' and control group	Reading Three weeks/once a week	(1) Sentence completion (2) Reading comprehension (3) Strategy use	Yes	No	<i>Positive</i> Generative précising resulted in better comprehension
Meskill, 1991 RCT Article	Language learning strategies advice: a study on effects of online messaging	USA ESL Adult N = 34	On-screen messaging advice on learning strategies	Proficiency	(1) Performance (2) Perceptions	No	No	<i>Negative</i> Qualitative data suggest positive effect
Najar, 1997 RCT (cluster) PhD	The effect of notetaking strategy instruction on comprehension in ESL texts	Japan EFL University 135/338 in 10 groups	Notetaking: (1) Awareness (2) Strategy	Reading Nine weeks	(1) Reading comprehension	Yes	Yes Two weeks	<i>Positive</i> Notetaking training group produced better reading comprehension results – but possibly due to learners doing more work at home.

Author, year, method, publication type	Study title	Setting, population (N = intervention/ total)	Strategy training intervention	Domain, intervention duration	Outcomes measured	Strategy change measured?	Delayed testing?	Results as reported
O'Malley <i>et al.</i> , 1985 RCT Article	Language strategy applications with students of ESL	USA ESL High school N = 75 (?) in three groups	(1) Metacognitive/ cognitive/socio-affective (2) Cognitive and socio-affective (3) Control	Listening Speaking (and Vocabulary) 8 days/50mins	Listening and speaking test performance	Observation but only of experimental groups	No	<i>Positive, negative and missing</i> Effective for speaking No effect found for listening Effect on vocabulary not reported
Ozeki, 2000 Comparative study PhD	Listening strategy instruction for female EFL college students in Japan	Japan EFL College – female N = 25/45 in experimental and control groups	Package of metacognitive, cognitive and affective strategies: directed and selective attention, self-evaluation, notetaking, inferencing, co-operation, etc.	Listening Six or seven months at 90 minutes per week	(1) Listening comprehension (2) Use of strategies	Yes	No	<i>Positive</i> Strategies package effective for listening and learners had positive attitude towards the training
Paulauskas, 1994 RCT PhD	The effects of strategy training on the aural comprehension of L2 adult learners at the high beginning/low intermediate proficiency level	Canada ESL Adult N = 51 in two intervention and one control groups	Predicting text content, summarising main ideas, questioning for comprehension, clarifying comprehension difficulties	Listening Four weeks at 3x1 hour	(1) Listening comprehension	Yes	Yes	<i>Positive</i> No differences between two interventions but both were effective for listening comprehension

Author, year, method, publication type	Study title	Setting, population (N = intervention/ total)	Strategy training intervention	Domain, intervention duration	Outcomes measured	Strategy change measured?	Delayed testing?	Results as reported
Raymond, 1993 RCT Article	The effects of structure strategy training on the recall of expository prose for university students reading French as a second language	Canada French University N = 43	Text structure strategy training (five top level strategies)	Reading 5 hours	Difference in recall of content, pre- and post-intervention	No	One month	<i>Positive</i> Intervention group scored higher on one of the test texts, but only on delayed test (no immediate post-test).
Sengupta, 2000 Comparative study Article	An investigation into the effects of revision strategy instruction on L2 secondary school learners	Hong Kong English Secondary school N = 78/108 in two intervention groups and one comparison group	Redrafting/ revision of first draft	Writing 12 essays over a year	(1) Attitude to writing (2) Gain score on writing task	No	No	<i>Positive</i> Redrafting is an effective strategy, but learners appeared to prefer traditional methods.
Seo, 2000 Comparative study PhD	Intervening in tertiary students' strategic listening in Japanese as a foreign language	Australia Japanese University N = 10 (self-assigned into two groups: intervention and control)	Cognitive and metacognitive strategies	Listening 19 weeks	Proficiency: comprehension of TV broadcasts	No	No	<i>Positive</i> Comprehension of broadcasts improved for intervention group, but author reports familiarity with test format may be a confounding influence.

Author, year, method, publication type	Study title	Setting, population (N = intervention/ total)	Strategy training intervention	Domain, intervention duration	Outcomes measured	Strategy change measured?	Delayed testing?	Results as reported
Talbot, 1995 RCT PhD	Metacognitive strategy training for reading: developing second language learners' awareness of expository text patterns	Hong Kong English (ESL) Higher Education/ tertiary N = 183/244 in three intervention groups and one control group	Metacognitive strategy awareness raising for text structures	Reading Five weeks at 10 x 60 minutes and two tests	Discussing incomplete texts, identifying and correcting illogical texts, identifying discourse patterns, unscrambling text (proxies for comprehension)	No (but partially yes)	No (partial)	<i>Positive</i> Training in text structure awareness was effective.
Thompson and Rubin, 1996 RCT Article	Can strategy instruction improve listening comprehension?	USA Russian University N = 24/36 in 2 intervention and 1 control groups	Metacognitive strategies: predicting content, listening for the known, listening for redundancy	Listening 5 weeks at 3x50 minutes per week	(1) Comprehension of video (2) Audio comprehension	No	No	<i>Mixed</i> Training showed positive effect on video test.

Appendix 4.2: Weight of evidence components for studies included in in-depth review (N = 25)

Item	A	B	C	D
	Reliability	Appropriateness	Relevance	Weight of evidence
Ayaduray and Jacobs (1997) Can learner strategy instruction succeed? The case of higher order questions and elaborated responses	Medium	Medium	Medium	Medium
Bimmel <i>et al.</i> (2001) Effects of strategy training on reading comprehension in first and foreign language	Low	Medium	Medium	Low
Bishop (2001) Using quality and accuracy ratings to quantify the value added of a dictionary skills training course	Medium	High	High	Medium
Burgos-Kohler (1991) The effect of a selected group of language learning strategies upon language development	Medium	High	High	Medium
Cadierno-Lopez (1992) Explicit instruction in grammar: a comparison of input based and output based instruction in second language acquisition	Medium	Medium	Medium	Medium
Carrell <i>et al.</i> , (1989) Metacognitive strategy training for ESL reading	Low	Medium	High	Low
EI-Koumy (1999) Effects of three semantic mapping strategies on EFL students' reading comprehension	Medium	Medium	Medium	Medium
Feyten <i>et al.</i> (1999) Consciousness raising and strategy use	Low	Medium	High	Low
Flaitz and Feyten (1996) A two phase study involving consciousness raising and strategy use for foreign language learners	Medium	Medium	High	Medium
Holunga (1994) The effect of metacognitive strategy training with verbalization on the oral accuracy of adult second language learners	High	High	Low	Medium
Kitajima (1997) Referential strategy training for second language reading comprehension of Japanese texts	Low	Medium	High	Low
Kusiak (2001) The effect of metacognitive strategy training on reading comprehension and metacognitive knowledge	Medium	Medium	High	Medium
Lawson and Hogben (1998) Learning and recall of foreign-language vocabulary: effects of a keyword strategy for immediate and delayed recall	Medium	Medium	High	Medium

Item	A	B	C	D
	Reliability	Appropriateness	Relevance	Weight of evidence
McGruddy (1995) The effect of listening comprehension strategy training with advanced level ESL students with students of English as a second language	Low	Medium	Medium	Low
McGuire (1999) Generative Précising as a Reading Comprehension Strategy for Adult ESL Learners	Low	Medium	High	Low
Meskill (1991) Language learning strategies advice: a study on the effects of on-line messaging	Low	Low	Low	Low
Najar (1997) The effect of notetaking strategy instruction on comprehension in ESL texts	Medium	High	High	Medium
O'Malley <i>et al.</i> (1985) Learning strategy applications with students of English as a second language	Medium	Low	High	Medium
Ozeki (2000) Listening strategy instruction for female EFL college students in Japan	Low	Medium	Medium	Low
Paulauskas (1994) The effects of strategy training on the aural comprehension of L2 adult learners at the high beginning/low intermediate proficiency level	Medium	High	Medium	Medium
Raymond (1993) The effects of structure strategy training on the recall of expository prose for university students reading French as a second language	High	High	High	High
Sengupta (2000) An investigation into the effects of revision strategy instruction on L2 secondary school learners	Medium	Medium	High	Medium
Seo (2000) Intervening in tertiary students' strategic listening in Japanese as a foreign language	Low	Medium	Medium	Low
Talbot (1995) Metacognitive strategy training for reading: Developing second language learners' awareness of expository text patterns	High	High	High	High
Thompson and Rubin (1996) Can strategy instruction improve listening comprehension?	Medium	High	Medium	Medium