Occasional Paper No. 2
October 2002

A Critical Review of Evaluations of Interprofessional Education

The Higher Education Academy
Health Sciences and Practice

www.health.heacademy.ac.uk

Published by:

Higher Education Academy
Health Sciences and Practice Network
Room 3.12 Waterloo Bridge Wing
Franklin-Wilkins Building
King's College London
150 Stamford Street
London SE1 9NH

ISBN No. 0-9542440-1-X

Correspondence to:

Dr Della Freeth
Reader in Education for Health Care Practice
Health Care Education Development Unit
City University
20 Bartholomew Close
London EC1A 7QN

Email d.s.freeth@city.ac.uk

A Critical Review of Evaluations of Interprofessional Education

This review was commissioned by the Learning and Teaching Support Network Health Sciences and Practice from the Interprofessional Education Joint Evaluation Team and published in May 2002.

The Interprofessional Education Joint Evaluation Team

Della Freeth
Marilyn Hammick
Ivan Koppel
Scott Reeves
Hugh Barr

CAIPE
UK Centre for the Advancement of Interprofessional Education

FOREWORD

This **'critical review of evaluations of interprofessional education'** is the second of a series of occasional papers commissioned by the Learning and Teaching Support Network (LTSN) Centre for Health Sciences and Practice. The first occasional paper by Professor Hugh Barr 'Interprofessional Education Today, Yesterday and Tomorrow' set the issue in historical and current context as well as suggesting future directions. This second paper develops the topic by reviewing systematically the literature on evaluations of interprofessional education. The choice of topic was based on a needs analysis of the Health Sciences and Practice community conducted at an early stage of the creation of the LTSN in 2000. One of the recurrent themes was that of interprofessional education.

The role of the LTSN UK national network is to promote good practices in Learning and Teaching in Institutions of Higher Education. To do so we first need to establish what is known about current practices in a variety of aspects of learning and teaching and to provide easy access to that literature which could help inform the community about existing evidence concerning what works and in what contexts. This implies that practices have been fully evaluated, which is not always the case.

This critical review of evaluations is written by Dr. Della Freeth, Dr Marilyn Hammick, Dr. Ivan Koppel, Scott Reeves, and Professor Hugh Barr, an interprofessional research group from the City, Oxford Brookes and Westminster Universities, and comprising radiography, social work, medicine, sociology and education. All are members of the UK Centre for the Advancement of Interprofessional Education (CAIPE).

Their findings form a solid basis and impetus for evidence-based practice and also for further evaluations. The authors found that studies were mostly focused on post-registration continuing professional development, in hospital or community based service delivery settings rather than in universities. The data is largely North American and most frequently represents nursing and medicine, followed by social workers, undifferentiated professions allied to medicine, pharmacists, physiotherapists and occupational therapists. Other disciplines are hardly represented despite the fact that the Health Services associated with the Centre encompass 30 or so subjects. This indicates that more good quality evaluations are needed within the UK, also at pre-registration level, and within a wider range of subject areas. The paper also critically reviews the quality of the evaluations and provides helpful suggestions for future studies to provide useful information for the development of good practice. It is certain to stimulate reflection and action.

Professor Catherine Geissler
Director LTSN Centre for Health Sciences and Practice

Contents

Tables

Figures

Boxes

1 Summary

This critical review of evaluations of interprofessional education, based on a systematic review of Medline, CINAHL and the British Education Index has revealed much about the nature and outcomes of interprofessional education, the evaluation of interprofessional education, and the dissemination of findings from evaluations of interprofessional education.

The review was conducted by a multidisciplinary team (section 4.2) that engaged in a lengthy, reflexive and iterative process of discussion, challenge, piloting, checking and modification to create shared understandings and rigorous procedures (chapter 5).

The studies we found mostly focused on post-registration continuing professional development. The interprofessional education usually occurred within the workplace or an employer's training facilities. Fewer than 30% of studies included pre-registration students, and the location of their interprofessional education was often a service delivery setting rather than the university. The post-registration interprofessional education could be subdivided into traditional staff development based on, for example, workshops and short-courses, or on the other hand, interprofessional education that occurred as a by-product of a quality improvement initiative. Thus we see interprofessional education and collaboration promulgated as a cause in its own right and as a problem-solving strategy.

The quality of studies was quite variable and we found it useful to focus on the higher quality studies from chapter 7 onwards. This data set is largely North American, evenly divided between hospital and community settings (with a highly correlated division of focus between acute and chronic conditions). The learning experience was almost always formal (e.g. workshop, seminar), although not usually overtly underpinned by any particular educational theory, and generally of medium or long duration. Nursing and medicine were the most frequently represented professions, reflecting the size and role diversity of these professions relative to others in health and social care.

Within the focused set of higher quality studies the evaluation designs were dominated by variants of before-and-after studies and longitudinal studies. These do seem to be the most profitable way of investigating processes and outcomes associated with interprofessional

education. However, convincingly demonstrating cause and effect remains a problem for many studies. Few studies addressed the longevity of any changes detected. More prospective studies and more longitudinal studies are required.

Studies based on quantitative data and limited interpretation dominated. We now need more interpretive and critical studies. Although expensive and relatively difficult to publish, there is much to be gained from well-conducted qualitative studies. Since most interprofessional education initiatives are multi-faceted, a greater number of mixed methods studies would be advantageous. It is comparatively difficult to secure funding for qualitative or mixed methods research studies of educational interventions for health and social care. There is much to be gained from addressing the multifaceted resistances to such studies.

The higher quality studies provided evidence of reasonable design in relation to evaluation questions, appropriate conduct and adequate analysis. They were often weaker in considering ethics, bias and the significance of their findings (statistical, practical, policy, etc.).

We grouped the reported outcomes of interprofessional education into six categories: learners' reactions, changes in attitude or perception, changes in knowledge or skill, behavioural changes, changes in the organisation or delivery of care, benefit to patients or clients. Most studies reported outcomes at more than one level. Studies evaluating university-based interprofessional education tended to focus on learners' reactions, attitudes and perceptions, knowledge and skill. This is not surprising. These learning experiences are about consciousness raising, preparation for future practice: a contribution to professional socialisation, an investment in the future. On the other hand, studies evaluating interprofessional education that occurred to augment a quality improvement initiative tended to focus on behavioural change, organisational change and patient benefit. This reflects their focus on specific problem solving in practice, *here* and *now*.

A glance through the reference section will reveal tremendous variety in the journals publishing evaluations of interprofessional education. These studies are often located with other work relating to the client group in question, or within the quality-improvement literature.

2 Acknowledgements

This critical review of evaluations of interprofessional education was commissioned by the Learning and Teaching Support Network for Health Sciences and Practice (LTSN HS&P). We are grateful to LTSN HS&P for funds to enable us, on this occasion, to focus on the professions within its remit. We are also grateful to our Universities (City, Oxford Brookes and Westminster) who support our ongoing systematic review work.

3 Correspondence

Dr Della Freeth: d.s.freeth@city.ac.uk

Dr Marilyn Hammick: m.r.hammick@bham.co.uk

Dr Ivan Koppel: ivan@ikop.demon.co.uk

Scott Reeves: S.Reeves-1@city.ac.uk

Professor Hugh Barr: barrh@westminster.ac.uk

4 Introduction

This report will not discuss the policy context of interprofessional education, nor outline its history. The interested reader is directed to a companion review, *Interprofessional Education: Today, Yesterday and Tomorrow* (Barr, 2002). Our intention is to make a contribution to the discourse about what kinds of interprofessional education should be encouraged and how best to evaluate the impact of these.

We are engaged in a systematic review process investigating studies that evaluate interprofessional education. This review reports part of our work.[a] It is a critical review of evaluations of interprofessional education involving one or more of the professions served by the Learning and Teaching Support Network for Health Sciences and Practice (LTSN HS&P – see section 5.2). Section 4.1 outlines the structure and content of the report.

Interprofessional education is widely seen as a way to develop collaborative practice among health and social care professions. The proposition that learning together may help people to work together more effectively is intuitively reasonable. However, this immediately leads to more difficult questions about: -

- What kind of interprofessional learning experiences one should aim for?
- What the outcomes of interprofessional education are?
- How the impact of interprofessional education can be detected?

This critical review will begin to illuminate these complex questions.

[a] In addition, from our Medline search, we have completed a separate analysis of evaluations of interprofessional education in which doctors or medical students participate. This will be published elsewhere.

4.1 The critical review report

Initially this report describes the process of identifying the literature for critical appraisal, including eligibility criteria, selection of abstracts and full papers, and the quality assurance processes that seek to enhance the reliability and validity of our findings. We also discuss the development of an outcomes model that allows the outcomes of interprofessional education to be classified across a number of equally important categories, *viz:*

- learners' reactions,
- modification of learners' attitudes or perceptions,
- acquisition of knowledge or skills,
- behavioural changes within professional practice,
- changes in the organisation and delivery of care,
- improvements in health or well-being of patients/clients.

It is interesting to note that we made provision for the emergence of outcomes that could not be accommodated within this typology with a category 'other'. After categorisation pf the reported outcomes in 217 evaluations of interprofessional education, this category remains empty.

The report also details results of the searches and provides a commentary on the process of data abstraction and recording. In section 5.4 we discuss our approach to evaluating the quality of the evaluations eligible for review. Finally, we discuss the limitations of the research presented herein (section 5.5).

The review findings are reported under two major headings to give firstly breadth and secondly depth. In chapter 6 (the broad picture) we report findings from the whole data set, i.e. including all eligible evaluations (n = 217). In particular, the location and type of educational opportunity and paradigmatic features of the evaluation are discussed. This is followed in chapter 7 (the focused picture) by a commentary on findings from a smaller number of higher quality studies (n = 53). In the next three chapters this is extended in respect of three aspects: -

- Characteristics of interprofessional education (Chapter 8).
- Features of evaluation design (Chapter 9).

- Reported outcomes (Chapter10).

Each aspect is subject to critical analysis within the context of the findings. Where possible further illustration is given by the use of specific citations, examples or more comprehensively, by vignettes selected from the studies included in the review. Our intention is to make a contribution to the discourse about what kinds of interprofessional education should be encouraged and how best to evaluate the impact of these.

Finally, in chapter 11, we draw some conclusions from our work in respect of interprofessional education, the evaluation of its impact and the dissemination of findings from these evaluations.

4.2 The review team

We are an interprofessional research group (radiography, social work, medicine, sociology and education). Each of us has experience in developing, delivering and evaluating interprofessional education. This has informed our contributions to formulation of policy in the field. We are all members of CAIPE (the UK Centre for the Advancement of Interprofessional Education). One guiding principal in our engagement with interprofessional education is the maintenance of critical reflexivity. We continually examine our own practice, in addition to theories and policies within the field.

We have been reviewing evaluations of interprofessional education since 1997. Four of us (HB, MH, IK, SR) contributed to a Cochrane review under the auspices of the Effective Practice and Organisation of Care (EPOC) subgroup (Zwarenstein, et al. 1999, Zwarenstein et al. 2001). This work provided excellent training in the process of systematic review, but was unsatisfactory in limiting the scope of the investigation to a narrow range of evaluation methodologies and a narrow range of outcomes. The current authors used this valuable learning to inform a series of more inclusive, but still systematic and rigorous, reviews of studies that evaluate interprofessional education (Koppel et al. 2001, Barr et al. 1999a, Barr et al. 1999b & Freeth et al. forthcoming).

5 Review Methodology

This review draws on evaluations of interprofessional education data collected in a systematic search of the international literature. The search strategy used is attached (Appendix I). Bibliographic databases searched were: -

- Medline 1966-2000: yield = 3374 abstracts
- Cumulative Index to Nursing and Allied Health Literature (CINAHL) 1982-June 2001: yield = 3054 abstracts
- The British Education Index (BEI), 1964-June 2001: yield = 49 abstracts.

> NB It is essential for systematic review teams to establish efficient error reducing systems for handling the large and complex data sets that result from comprehensive searches. Software packages that create databases of bibliographic references are very helpful, particularly with the filtering that is necessary to identify duplicated references (usually arising from overlap between bibliographic databases). We used *Reference Manager* software and one member of the team (IK) took lead responsibility for searching, then checking the database of abstracts. More details of our error reducing and quality strategies can be found in Reeves *et al* (2002). Aspects of the quality strategies are also discussed in sections 5.1, 5.3 and 5.4.

5.1 Selection of eligible evaluations

A two-stage process was employed in the selection of studies eligible for the review. Initially, each abstract obtained from the searches was evaluated by at least two members of the review team, asking: -

- Does this study describe interprofessional education?
- Has the education been evaluated?

In answering the first question we were guided by a negotiated definition of interprofessional education that required: -

> **Members (or students) of two or more professions associated with health or social care, to be engaged in learning *with, from* and *about* each other**.

Eligibility required an interactive element to the learning. Evaluations of education initiatives where members of several professions, for example, received shared lectures (i.e. shared listening, not interactive learning), did not qualify for inclusion.

In assessing whether the study had been evaluated we included evaluation methodologies that went beyond those permitted for the EPOC Review. Thus we included not only randomised controlled trials, interrupted time series, and controlled before and after studies, but also qualitative studies, quasi-experimental studies with or without controls, cost-benefit analyses and so forth.

Where both questions were answered positively full papers were obtained. This process generated 309 references from Medline, 103 new additions from CINAHL, then 5 new additions from the BEI. Full papers were obtained for these 417 studies.

Each full paper was scrutinised in the same double blind manner. If the two reviewers agreed that the paper reported an evaluation of interprofessional education, it was passed to SR for abstraction, coding and entering into a Statistical Package for the Social Sciences (SPSS) data file. As a quality check, other members of the team additionally coded approximately 10% of eligible studies. At both the abstract and full paper review stage discrepancies between the judgements of different review team members were resolved through discussion and then, if necessary, by referral to the whole team.

The review was purposefully inclusive, in respect of the definition of education and evaluation methodologies. For example, an interprofessional workshop convened as part of a quality improvement initiative to address identified weakness within service delivery, would be included even if the authors of that study did not identify this activity as interprofessional education (because of their focus on quality improvement).

We also took an inclusive approach to the outcomes of the interprofessional education. Those affecting participating learners, service users and organisations were all considered. Early in our systematic review work, and in common with other evaluators, we recognised the value of Kirkpatrick's model for summarising the outcomes of educational outcomes Figure 1 (see Thackwray (1997) for a full discussion of this). Over time our work has led to a modified form of Kirkpatrick's four-level model of educational evaluation. Through an iterative approach, informed by the evaluations we have studied, we have added two further levels (Figure 2).

Figure 1: Kirkpatrick's Model of Educational Outcomes

1	Reaction	To the educational experience
2	Learning	Mainly conceptualised as the acquisition of skills and knowledge
3	Behaviour change	Focused on whether participants now do things differently and the application of the learning to practice.
4	Results	Particularly in relation to intended outcomes.

Three guiding principals of the original Kirkpatrick model also underpin our model, namely that: -

- Outcomes in each of the areas are not hierarchical.
- The aim is to encourage more holistic and comprehensive evaluations to better inform future policy and development.
- There is acknowledgement that at each level it becomes progressively more difficult to gather trustworthy data related to the educational intervention.

Figure 2: Our Model of Outcomes of Interprofessional Education

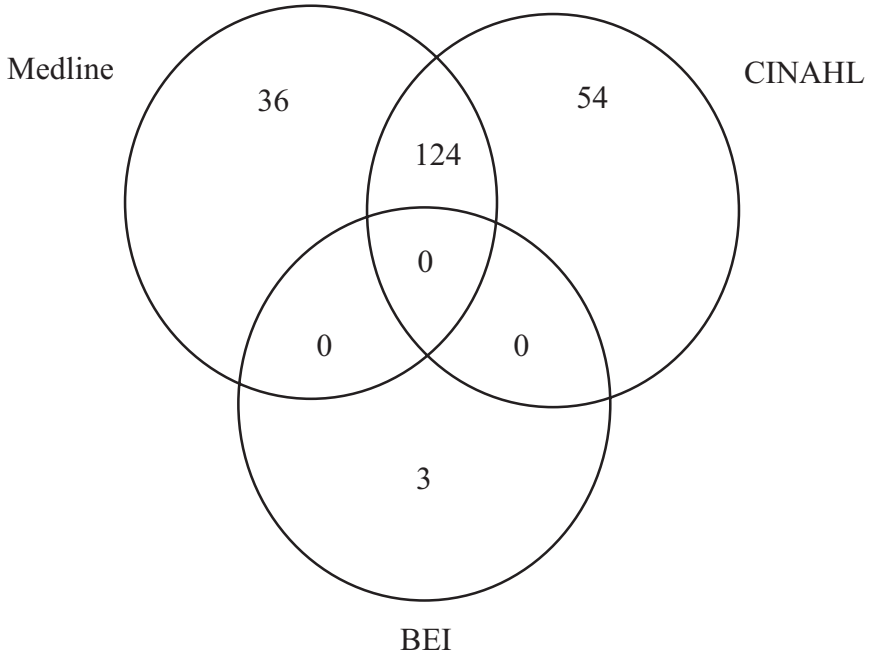
1	Reaction	Learners' views on the learning experience and its interprofessional nature.
2a	Modification of attitudes / perceptions	Changes in reciprocal attitudes or perceptions between participant groups. Changes in perception or attitude towards the value and/or use of team approaches to caring for a specific client group.
2b	Acquisition of knowledge / skills	Including knowledge and skills linked to interprofessional collaboration.
3	Behavioural change	Identifies individuals' transfer of interprofessional learning to their practice setting and changed professional practice.
4a	Change in organisational practice	Wider changes in the organisation and delivery of care.
4b	Benefits to patients / clients	Improvements in health or well being of patients / clients.

5.2 Current yield of evaluations

Medline yielded most abstracts because of the number and range of journals it reports. Therefore, the two-stage selection process began with abstracts identified by that database. This resulted in a data set of 162 eligible evaluations of interprofessional education. Checking these studies against the CINAHL bibliographic database showed that 124 could also be located within CINAHL. The remaining CINAHL abstracts were then processed as described above and this yielded a further 55 evaluations, bringing the total to 217 evaluations of interprofessional education. Three of the full papers obtained from the BEI search met our inclusion criteria, bringing the total to 220.

The 220 evaluations of interprofessional education were further scrutinised for participation by one or more of the groups served by the Learning and Teaching Support Network for Health Sciences and Practice.[b] This reduced the data set to 217 eligible evaluations, distributed between the bibliographic databases as shown in Figure 3. The analysis of data from these 217 studies[c] informs this critical review. If CINAHL alone had been searched 39 evaluations of interprofessional education involving professions served by the LTSN HS&P would have been missed.

Figure 3: Venn diagram illustrating the bibliographic source of the 217 studies informing this critical review



[b] The LTSN HS&P lead subjects are: Audiology, Chiropdy & Podietry, Chiropractic, Complementary Medicine, Environmental Health, Gerontology, Health Promotion & Education, Health Visiting, Healthcare ethics, Medical Engineering, Medical Physics, Midwifery, Nursing, Nutrition & Dietetics, Occupational Therapy, Optometry, Orthoptics, Osteopathy, Pharmacy, Physiotherapy, Prosthetics and Orthodontics, Public Health, Radiography & Radiotherapy, Rehabilitation Sciences, Speech & Language Therapy.
NB The studies reported in this review, whilst restricted to those including one or more of the professional groups represented by the LTSN HS&P lead subjects, also included participants from professions of interest to other LTSN subject centres, particularly medicine, dentistry and social work (see section 8.1 for a fuller description of professional mix).

[c] "Studies" are almost equivalent to publications. Just four papers/reports described multiple evaluations that needed to be coded separately: Chessman *et al* (1996) – two studies; Headrick *et al* (1996) – four studies; Mitchell *et al* (1996) – three studies; Stanford & Yelloly (1994) – two studies. This will be discussed further in chapter 7.

5.3 Data abstraction and coding

Data abstraction involved the manual completion of a validated abstraction sheet for each study. This allowed a summary of the content of each paper. Data were collected on the context and rationale for the interprofessional education, its aims, content, duration, pattern of delivery, learning and teaching methods, and who participated. All reported outcomes were categorised. Characteristics of the evaluation were coded including: rationale, methods, analytical approach, and indicators of trustworthiness and ethical conduct.

Two data abstraction sheets were used for this process: one for quantitative and the other for qualitative studies (Appendix 2). Both were used to capture data from mixed methods studies. The evolution of the data abstraction sheets was informed by the original abstraction sheet accepted for the EPOC review, our research and evaluation experience, research methods texts and evaluative criteria suggested by others. The quantitative data abstraction sheet at version six is at steady state. It is more highly developed due to the number of studies using this data collection method. The qualitative sheet at version two remains a prototype, due to the limited number of studies analysing this type of data.

Each member of the review team, using the same small set of studies, tested prototype abstraction sheets. Difficulties and differences were discussed and resolved, sometimes leading to modification of the data abstraction sheet. This iterative process is lengthy and it makes an important contribution to the rigour of the review. The debate hones shared definitions of the phenomena being studied and this, in turn, determines the quality, clarity and comprehensiveness of data recording.

All the data from the abstraction sheets that could be coded as categorical (nominal) variables, plus a small number of numeric variables, were transferred to a data file that could be read by the statistical software SPSS. This permitted easy partitioning and exploration of the data set. Any relationships that seemed to exist were checked through the data extraction sheets and original papers. Time was allowed for *cleaning* the data file (that is, searching it for data entry errors and omissions) as failure to do this may compromise subsequent analyses and inferences.

5.4 Evaluating quality

Preliminary reading of the 217 studies in this review indicated that they were variable in quality. A mechanism for identifying the most trustworthy and illuminating studies was required. Two variables were added to our data abstraction sheets to grade quality: one evaluating the quality of the study design in relation to its research questions, and another evaluating the quality of the information provided in the published account.

The quality of the study was judged on a five-point scale (5 = best score) that took account of the appropriateness of the evaluation design in relation to its aims and the questions addressed. Judgements were made about whether the number of participants was appropriate to the design, whether selection of participants rested on clear criteria, whether bias, validity and reliability or authenticity and trustworthiness had been considered, and other criteria relevant to the evaluation design. Thus, where the research questions lent themselves to a controlled study utilising quantitative data, a well conducted randomised controlled trial, interrupted time series study, or controlled before-and-after study would score five, provided sampling, analysis and consideration of ethics and bias (for example, loss to follow-up) were appropriate. Similarly, where the research questions were process orientated, well conducted ethnographic studies would score five, provided the issues of reflexivity, ethics, relevance, authenticity, and trustworthiness were appropriately considered, in addition to clarity regarding inclusion/exclusion from the study and evidence of rigorous analysis. Other evaluation designs that would attract a high score would be rigorous action research and well conducted longitudinal studies.

Evaluation designs that were awarded mid-range scores included competently conducted uncontrolled before-and-after studies, and observational studies that were systematic, with clear objectives and inclusion/exclusion criteria, but, for example, were lacking in their depth of analysis or attention to reflexivity. Evaluation designs that scored only one included post-intervention studies and observational studies that lacked the clarity of objectives and critical thinking.

The quality of the evaluation had to be inferred from the published accounts. This means that a high quality study might be evaluated poorly because the information provided in the published account was insufficient for readers to evaluate the quality of design, conduct and analysis. A separate score was allocated to each study to reflect the quality of the written information. This was also on a five-point scale (5 = highest). Those studies scoring highly provided a clear account of the context for the interprofessional education and its evaluation, a clear rationale for the evaluation, stated the evaluation questions, provided sufficient information in relation to sampling, ethics, bias, reliability, validity, authenticity, trustworthiness (as appropriate to the design), described the analysis in sufficient detail, presented results and inferences clearly, considered the significance (statistical and practical) of findings, and made appropriate recommendations.

It was the norm for studies to address only some of these issues and in varying degrees of depth. This to be expected when authors write an account prepared for a particular audience, meeting the publication requirements of a particular journal. This effect was most apparent in studies reported within tight word limits. In these cases the emphasis was placed upon context, design and findings, generally leaving the reader somewhat uncertain with respect to rationale, conduct, analysis and the basis for inferences.

The distribution of quality scores awarded in this review is displayed in Table 1. Clearly, these scores are correlated (Spearman's Rank Correlation Coefficient = 0.622). This is to be expected since it is difficult to evaluate the quality of a poorly described study (we will return to this point in chapter 9). The absence of any top scoring studies may reflect unrealistically high ideals on the part of the review team. However, it is fair to say that within the extensive searches we have conducted, there are very few excellent evaluations of interprofessional education. Some evaluations did attract high scores, with seven scoring four for both design and information quality.

All the studies help to further the understanding of how interprofessional education is evaluated, what kinds of outcomes it produces and what links can be found between the various features of the education programme. However, for this critical review we have focused upon the more robust and more clearly reported evaluations. Thus, subsequent analyses eliminated studies that scored 1 or 2 on either of the quality variables shown in

Table 1. This exclusion criterion reduces the data set to 53 evaluations of interprofessional education indicated in the shaded portion of Table 1. This approach allowed us to look at the higher scoring studies in greater detail and draw firmer conclusions.

		Quality of study				
	Score	**1**	**2**	**3**	**4**	**5**
Quality of information	**1**	39	22	3	0	0
	2	14	41	12	1	0
	3	4	24	34	5	0
	4	1	3	7	7	0
	5	0	0	0	0	0

Table 1: Cross-tabulation of Scores for Study and Information Quality

5.5 Review limitations

Earlier in this report we commented on the characteristics of the review team and, in particular, the position we take with respect to our own practice, including our research practice. We are committed to examining the limitations of our work and commenting on how these shape our findings and conclusions. This is done in the spirit of sharing challenges we encountered during our work for this review. It highlights the influences on the data we have used. We hope that this provides an insight for the community of teachers and practitioners about what can be learnt from a systematic search of the literature and a critical review of the results. The following section is not an exhaustive list of limitations and their influences; rather it seeks to outline the major boundaries within which this critical review was written.

The use of bibliographic databases to identify the potential literature for our review provided an efficient but not foolproof source of material. A complex search strategy in a field such as ours, where the terminology is still inconsistent across international and professional borders,

had fault lines that gave rise to considerable challenges for the information scientist who provided initial advice for this aspect of our work. These challenges remain for the team members who lead on adapting and updating the search strategy to cope with the idiosyncrasies of different bibliographic databases and different user interfaces. We recognise that this may have influenced the final data set.

In addition, we highlighted earlier the growth in publications in the last decade. This is likely to have continued and our data set is bounded by the date of the latest search. Resources permitting, regular updates of critical reviews are important.

Our source material has been influenced by the bibliographic databases we have been able to search (Medline, CINAHL and the BEI) and the journals that these databases abstract. The geographic distribution of the journals abstracted by these databases has inevitably accentuated the North American bias and English language bias within studies selected for this review. There may also have been a subtler privileging of quantitative studies. Returning to language bias, the review team was able to evaluate abstracts published in English and French. However, we did not have resources for translation from other languages. In practice this meant 15 abstracts were discarded from the 6475 that were obtained from the bibliographic searches.

Resource constraints also prevented us from fully utilising other potential sources of interprofessional education evaluations. These would have included references cited in eligible full papers, the grey literature, web-based reports and papers in relevant journals not yet abstracted by Medline, CINAHL or the BEI. Nevertheless, members of the team have previously undertaken searches of these sources for reviews of the position of interprofessional education within the United Kingdom (Barr & Shaw, 1995; Barr & Waterton 1996; Barr et al. 2000). This background information leads us to believe that an analysis of the data set we have obtained from our systematic reviews of Medline, CINAHL and the BEI, despite its limitations, is a useful contribution to knowledge and debate in the United Kingdom and beyond.

As is generally the case, there are two other sources of bias in this critical review. Firstly, without exception the eligible papers reported positive results, although some also addressed

setbacks and barriers. This is not unusual but it is important. Essentially it means that evidence of what hinders the effectiveness of interprofessional learning is in short supply. We wish to encourage researchers to share both negative and positive findings with their peers, and thus to contribute to all aspects of the interprofessional learning discourse.

Secondly, and importantly, although we seek to maintain critical reflexivity as individuals and as a research team, and have been as vigilant as possible about our quality control processes, we recognise the bias that we ourselves bring to this report. This applies to the selection of abstracts, full papers and data, to the analysis and synthesis of the findings and to the judgements we have made.

Finally, systematic bias is not the same as random error. The quality control mechanisms we devised were aimed minimising bias and error. Checking of processes and results has occurred at every stage of our work. We hope that errors are few and insignificant, but we would appreciate notification of any you notice. This work will be updated periodically, providing opportunities for errors to be corrected. In particular, if your evaluation of interprofessional education has slipped through the net of our electronic searches and subsequent checking of over 6000 references, please tell us about it.

6 The Broad Picture (all eligible studies)

6.1 Geography and chronology

Table 2 shows the distribution of year of publication for the included studies. The studies were located as follows: USA (170, 78%), UK (26, 12%), Canada (9, 4%), Australia (8, 4%), Malta (1, <1%), Norway (1), Saudi Arabia (1) and Turkey (1). Thus, most studies included in this critical review are North American and have been published since 1995.

Year	Frequency
1969-80	10 (5%)
1981-85	8 (4%)
1986-90	15 (7%)
1991-95	56 (26%)
96-2001	128 (59%)
Total	**217**

Table 2: Distribution of Year of Publication

6.2 Health and Social Service sector

The studies were reasonably evenly split between evaluations of interprofessional education related to hospital-based care and, on the other hand, related to community-based care (see Table 3).

Category	Frequency
Hospital-based	104 (48%)
Community-based	87 (40%)
Mixed	10 (5%)
Not clear	16 (7%)

Table 3: Care Sector Distribution

6.3 Career stage and location of interprofessional education

Most of the studies evaluated interprofessional education that occurred after initial professional qualification, for example, in-service training, continuing professional development (CPD) workshops or courses and postgraduate studies (Table 4). At pre-qualification level the location of the evaluated interprofessional education was more often a university than a service setting. However, 33 studies did report pre-qualification education that took place wholly or partly within service settings. At post-qualification level most interprofessional education is delivered in service settings or their associated staff training facilities.

Career stage	Frequency	Primary educational location	
Pre-qualification	55 (25%)	• Higher education institution • Service setting (e.g. clinical placement) • Mixed	29 11 15
Post-qualification	150 (69%)	• Higher education institution • Service setting • Mixed	13 134 3
Mixed	12 (6%)	• Higher education institution • Service setting • Mixed	5 3 4

Table 4: Career Stage and Location

The group of 134 studies concerning post-qualification studies in service settings can be further subdivided. Firstly, there were many quality improvement initiatives (91 studies, 68%) underpinned by the theories of continuous quality improvement (CQI), total quality management (TQM), the learning organisation, or clinical audit. These often focused on guideline development or improvement (84 studies). Secondly, the remaining 43 studies (32%) reported staff development activities such as workshops, short-courses, or problem-solving groups. These were underpinned by theories of adult learning and problem-based learning.

6.4 Type of interprofessional education

The education reported in each study was coded as formal, informal or mixed (Table 5). The formal category indicates learning opportunities such as modules within pre- or post-qualification courses, staff development workshops, seminar series, etc. The informal category included learning experiences such as interprofessional meetings to formulate best practice guidelines within a quality improvement initiative. Most published studies of interprofessional education evaluate formal learning opportunities.

Category	Frequency
Formal	142 (65%)
Mixed	54 (25%)
Informal	21 (10%)

Table 5: Formal or Informal Education

6.5 Evaluation design and analytical orientation

Table 6 and Table 7 show the distribution of evaluation designs and analytical orientations found within the included studies. It can be seen that the evaluation design was not discernable within the published accounts in almost 12% of studies. There is a large cluster (26%) of studies evaluating interprofessional education with the relatively weak post-intervention design (for example post-workshop questionnaire or interview). There is a preponderance of quantitative approaches (132 studies, 61%) although only 32 of these employed tests for statistical significance. The remainder were descriptive reports based on quantitative data. The analytical orientation was not discernable in 28% of studies. These observations raise concerns about the quality of some of the evaluations of interprofessional education within the data set: hence our focus in subsequent chapters on the 53 studies that scored three or higher for each of the quality variables described in section 5.4.

Design	Frequency
Post-intervention, single time point	56 (26%)
Post-intervention, single time point, with control group	6 (3%)
Post-intervention with follow-up	6 (3%)
During-and-after study	1 (<1%)
Before-and-after study	46 (21%)
Controlled before-and-after	8 (4%)
Before-during-and-after	6 (3%)
Before-and-after with follow-up	11 (5%)
Longitudinal	45 (20%)
Longitudinal with control group	1 (<1%)
Randomised control trial	2 (1%)
Action research	1 (<1%)
Case study	1 (<1%)
Not given	27 (13%)

Table 6: Evaluation Design

Data Analysis	Frequency
Quantitative	132 (61%)
Qualitative	7 (3%)
Mixed	17 (8%)
Not given / unclear	61 (28%)

Table 7: Analytical Orientation

7 The Focused Picture (higher quality subset)

As described in section 5.4, the variable quality of studies emerging from our searches made it advisable to grade evaluations: firstly for the strength of their design in relation to the reported evaluation questions and, secondly, for the quality of the information provided in the published account. Fifty-three studies scored three or higher on each of the five-point quality scales we developed. This and subsequent chapters focus on this more robust subset of evaluations of interprofessional education.

The 53 stronger studies, apart from two emerging from the BEI search, were evenly divided between Medline (42 evaluations) and CINAHL (42 evaluations), with 33 studies (79%) appearing in both these bibliographic databases. They reported interprofessional education in the USA 37 (70%), UK 11 (21%), Canada 3 (6%), Australia 1 (2%) and Norway 1 (2%). Thus UK studies form a higher proportion of the stronger studies data set than they form within the complete data set. The UK studies are Carpenter (1995), Crawford *et al* (1998), Freeth & Nicol, (1998), Long (1996), Midence (1991), Nash & Hoy (1993), Parsell *et al* (1998), Rutter & Hagart (1990), Stanford & Yelloly *et al* (1994)[d] and Stein & Brown (1995). Brief descriptions of these studies can be found in Box 1.

[d] Stanford & Yelloly *et al* (1994) reports the evaluation of two interprofessional courses; one lasting ten weeks and the other two years. It proved impossible to code this substantial research report as a single study. Therefore, each course appears within our database as a separate study.
One other paper within the higher quality subset needed to be split into separate studies: Mitchel *et al* (1996) 'Three teams improving thrombolytic therapy' (USA) was coded as three studies.

Box 1: The stronger United Kingdom studies

Carpenter (1995) reports a one-day interprofessional course for final year medical and fourth year undergraduate nursing students. The theoretical underpinning for the course was contact theory. It aimed to enhance students' attitudes and knowledge of each other's profession and to explore ways of working together in a co-operative, patient-focused fashion. Based in interprofessional pairs, 23 medical students and 16 nursing students worked together exploring and discussing issues around delivering effective patient care. The course was evaluated by pre/post questionnaires. Results indicated that the students enjoyed their shared learning experiences and felt their knowledge of interprofessional working had improved. However, attitude changes were mixed: while nursing students' attitudes toward the medical students had improved, the medical students' attitudes towards the nursing students were unchanged or slightly poorer.

Crawford et al (1998) present findings from their evaluation of a series of one-hour interprofessional workshops for doctors and nurses based in a hospital accident and emergency (A&E) department. The workshops aimed to improve the care of deliberate self-harm patients who present in the department. In total 45 nurses and 15 doctors attended one of the workshops held over a three-week period. An audit of patient notes undertaken before and after the sessions indicated that the A&E staff had completed notes in a more accurate and comprehensive fashion following the sessions. In addition, it was found that there was an increase interprofessional liaison between A&E staff and the hospital's parasuicide team. Despite this gain, it was found that the number of patients who went on to contact the parasuicide team remained unchanged.

Freeth & Nicol (1998) describe an evaluation of a pilot interprofessional programme for 7 final year medical students and 7 newly qualified nurses that aimed to develop a range of clinical and communication skills. The course comprised of four half-days, delivered over four weeks. During the course participants received interprofessional team teaching. Participants jointly worked on a number of patient scenarios, during which they were encouraged to learn from one another, share their knowledge and also discuss professional stereotypes and teamwork. Questionnaire, interview and observational data were collected. It was found that participants enjoyed their interprofessional learning and felt they had increased their knowledge of clinical and communication skills. It was also found that both facilitators and participants had markedly different teaching/learning styles. These appeared to be a result of their traditional professional socialisation.

Long (1996) reports findings from an evaluation of a two-day team-building workshop delivered to two primary health care teams. Each team consisted of a general practitioner, a practice nurse, a receptionist, a health visitor, plus a district nurse in one team and a midwife in the other. The purpose of these workshops was to improve interprofessional co-ordination and understanding between different professional groups through discussion and team problem solving. Team members were interviewed before and after the workshops. The research findings indicated that participants enjoyed their interprofessional learning and reported that they had a better awareness of each other's roles. In addition, they reported that communication had improved and their collective practice had changed as a result of the workshops.

Midence (1991) describes the development and evaluation of informal work-based interprofessional sessions for a team of two occupational therapists, a physiotherapist, a clinical psychologist, a speech therapist and six support staff based at a therapeutic day unit for clients with learning difficulties. The aim of these sessions was to increase the levels of staff-client engagement and interaction. The sessions, held on a weekly basis over a number of months, sought to enhance understanding of issues around client care and to allow team members to discuss their ideas for improving teamwork and service delivery. Baseline and five month follow-up observational and questionnaire data collected on levels of staff-client engagement and interaction revealed that a number of improvements were made.

Nash & Hoy (1993) describe the development and evaluation of a three-day residential workshop for general practitioners and district nurses focused on enhancing their understanding of aspects of delivering interprofessional terminal care. The authors present findings from 11 workshops they have delivered and evaluated, attended by 47 general practitioners and 47 district nurses. Data were collected via pre/post questionnaires. These revealed that participants valued the interprofessional learning and generally felt more confident about working together in this area. It was also found that participants' understanding of teamwork had improved. The authors go on to outline ideas for developing follow-up workshops to consolidate the initial successes of this learning experience.

Parsell et al (1998) examined the impact of a two-day pilot interprofessional course on 28 final year students from occupational therapy, orthoptics, radiotherapy, nursing, physiotherapy, medicine and dentistry. Working in small interprofessional groups, students discussed and engaged in problem-solving relating to teamwork and collaboration. The course was evaluated with pre/post questionnaires to assess changes in student attitudes and knowledge. To track the longer-term impact of the course, six week follow-up data were also collected.

Findings revealed that the course increased students' knowledge and understanding of teamwork and helped to develop more positive attitudes of other professional groups. The authors stressed the importance of interprofessional planning in the development of successful interprofessional courses.

Rutter & Hagart (1990) describe the development and evaluation of an interprofessional diploma in alcohol counselling and consultation. Participants included nine nurses, ten voluntary workers, four social workers, a probation officer, an occupational therapist and a general practitioner. The course was delivered in one-two week full-time blocks over a period of one year. To assess the impact of this course on participants' a controlled before-and-after research design was adopted with the collection of questionnaire, interview and observational data. Findings revealed that participants enjoyed the course and reported that their knowledge of issues around interprofessional collaboration within the field of alcohol counselling improved. However, it was found that their attitudes and the quality of their interactions over the duration of the course remained largely unaltered. Data from the 'control' participants also showed no change.

Stanford & Yelloly et al. (1994) report on a UK study that evaluated the impact of two interprofessional courses that aimed to enhance knowledge and inter-agency collaboration in the area of child protection. Two jointly validated (ENB & CCETSW) courses were evaluated by the project. The first was a short introductory course on issues around inter-agency child protection that lasted for 10 weeks. The second, longer course was offered over a two-year period. This course led to a master's degree. In total 26 participants took part in these courses (16 in the short course, 10 in the longer course). Unfortunately, it was not clear which of these participants completed both courses. Professions represented were nursing, health visiting, midwifery, police and youth workers. The evaluation adopted a case study approach, collecting both process (observations, documentary data) and outcome (interviews, questionnaire) data. Findings revealed a high level of satisfaction with both courses and improved knowledge around the contributions each profession makes to child protection.

Stein & Brown (1995) discuss the development, delivery and evaluation of a three-day interprofessional course for doctors, police officers and voluntary workers, social workers, nurses and occupational therapists, all working with adults with learning disabilities who have been abused. The course aimed to increase participants' understanding of the different roles and responsibilities of the professionals involved in this area and also to assist participants to work collaboratively to serve this client group. The course was delivered through a mixture of lectures and small group discussions. Questionnaire data, collected

before and after the course, revealed that participants felt they had acquired a firmer knowledge of the different groups involved in the care of adults with learning disabilities. The authors asserted that the course had led participants to engage in a collaborative mode of information sharing rather than the competing mode that can characterise interprofessional work in this area.

There is a strong North American bias to the higher quality studies (and the full data set). The findings and inferences of this review must be treated with care in this respect. They may not be applicable to other health care education systems. The North American bias is likely to be multifaceted. Influences include: the geographic source of journals abstracted by Medline and CINAHL, the relative size, in particular, of the US population and health care system, and relative strength of the culture of evaluation between the different countries.

The right-hand columns in Tables 8-13 show how the key descriptive variables, reported in the previous chapter for all eligible studies, were distributed among the higher quality studies. The focused picture of the higher quality subset is generally representative of the broader picture of the entire data set. However, there are proportionally fewer 'not clear' and 'not given' entries in the focused picture and only two studies span primary and secondary care sectors (Brown, 2000; Stanford & Yelloly *et al*, 1994) (Table 9).

Year	Broad	Focused
1969-80	10 (5%)	2 (4%)
1981-85	8 (4%)	2 (4%)
1986-90	15 (7%)	2 (4%)
1991-95	56 (26%)	21 (40%)
96-2001	128 (59%)	26 (49%)
Totals	**217**	**53**

Table 8: Distribution of Year of Publication: comparison of the broad and focused picture

Category	Broad	Focused
Hospital-based	104 (48%)	24 (45%)
Community-based	87 (40%)	26 (49%)
Mixed	10 (5%)	1 (2%)
Not clear	16 (7%)	2 (4%)

Table 9: Care Sector Distribution: comparison of the broad and focused picture

Several results collectively provide an indication of the profile of interprofessional education. Most usually, it was a formal educational intervention (Table 10), over several days or weeks, as part of a qualified practitioner's continuing professional development (Table 11) and linked into clinical audit or quality improvement processes (see section 8.5). Less often, the interprofessional education was for pre-registration students, formal, lasting less than seven days and unlikely to contribute towards an academic award (see section 8.2).

Category	Broad	Focused
Formal	142 (65%)	39 (74%)
Mixed	54 (25%)	11 (21%)
Informal	21 (10%)	3 (6%)

Table 10: Formal or Informal Education Programme: comparison of the broad and focused picture

Category	Broad	Focused
Pre-qualification	55 (25%)	13 (25%)
Post-qualification	150 (69%)	39 (74%)
Mixed	12 (6%)	1 (2%)

Table 11: Career Stage Distribution: comparison of the broad and focused picture

In the set of higher quality studies (the focused picture) there are fewer post-intervention studies and a greater proportion of before-and-after studies (Table 12) and mixed methods studies (Table 13).

Design	Broad	Focused
Post-intervention, single time point	56 (26%)	2 (4%)
Post-intervention, single time point, with control group	6 (3%)	1 (2%)
Post-intervention with follow-up	6 (3%)	1 (2%)
During-and-after study	1 (<1%)	1 (2%)
Before-and-after study	46 (21%)	24 (46%)
Controlled before-and-after	8 (4%)	6 (11%)
Before-during-and-after	6 (3%)	4 (8%)
Before-and-after with follow-up	11 (5%)	4 (8%)
Longitudinal	45 (20%)	9 (17%)
Longitudinal with control group	1 (<1%)	0
Randomised control trial	2 (1%)	1 (2%)
Action research	1 (<1%)	0
Case study	1 (<1%)	0
Not given	27 (13%)	0

Table 12: Evaluation Design: comparison of the broad and focused picture

Data Analysis	Broad	Focused
Quantitative	132 (61%)	36 (68%)
Qualitative	7 (3%)	1 (2%)
Mixed	17 (8%)	12 (23%)
Not given / unclear	61 (28%)	4 (8%)

Table 13: Analytical Orientation: comparison of the broad and focused picture

Tables 8-13, however, only show part of the character of the interprofessional education evaluated in the higher quality studies. The following chapters describe more fully the interprofessional education and its evaluation for those 53 studies.

8 The Character of the Interprofessional Education

8.1 Professional mix

The pattern of professional participation in the interprofessional learning we reviewed is shown in Table 14. The number of professional groups within the data set is perhaps smaller than might have been expected. No studies included representatives of audiology, podiatry, optometry or radiography, unless some of these groups were hidden in the studies under the group 'PAMs' (professions allied to medicine), which is used by some study authors. Apart from places where we are quoting study authors, we use the term Allied Health Professionals (AHPs) for this group. Publications where participants were members of the Allied Health Professions often lacked detail on who made up this diverse group. Authors could be more helpful in this respect in the future.

Profession	Number of studies
Nurses	49 (93%)
Doctors	43 (81%)
Social workers	21 (40%)
PAMs (undifferentiated by study authors)	18 (35%)
Pharmacists	10 (19%)
Physiotherapists	11 (21%)
Occupational therapists	9 (17%)
Psychologists	5 (9%)
Dentists	3 (6%)
Speech therapists	2 (4%)
Health Visitors and Midwives	2 (4%)
Dieticians	1 (2%)
Unspecified members of a multidisciplinary team	1 (2%)
Other (e.g. administrators, technicians, clergy)	23 (44%)

Table 14: Professions participating (higher quality studies)

Nurses and doctors are the two largest practitioner groups; they are likely to be more prevalent in health and social care teams. This makes it more likely that they will be participants in interprofessional education. Table 14 counters any stereotype that doctors are unwilling to participate in learning with those from other professions.

There is evidence of interprofessional learning across service sector boundaries, for example, when health professionals share their learning with social workers, police, the clergy or representatives of the voluntary sector (e.g. Stein & Brown, 1995). This is the type of interprofessional education advocated in the many policy documents that call for increased inter-agency collaboration.

There is a high degree of cultural specificity associated with certain roles and/or role titles. For example the role of Health Visitors is well developed in the UK but does not have a direct equivalent in the USA (from whence most of the included studies originate). Furthermore, health visitors may be under-reported as participants since, in the UK, these community-based practitioners all have a first level nursing qualification so they may have been recorded as 'nurses'.

Reviewers should also take note of the influence of culturally specific professional role terminology on the results of their searches. Such artefacts of classification can result in unforeseen bias in a review of the international literature.

8.2 Stage of education

Most of the studies (39, 74%) evaluated interprofessional education that contributed to continuing professional development (CPD), rather than initial qualification (Table 15). It was unusual for learning to contribute towards an academic award (only 3 studies, 6%): see for example (Rutter and Hagart 1990).

	Initial qualification	CPD	Mixed
Service-setting	6 (11%)	33 (62%)	0
Higher education only	4 (8%)	5 (9%)	1
Mixed	3 (6%)	1 (2%)	0

Table 15: Location and level of interprofessional education (higher quality studies)

8.3 Setting and patient/client condition

Most of the studies (39, 74%) described interprofessional education, whether pre- or post-registration, that was sited in service delivery settings (Table 15). It appears logical and feasible to encourage shared learning for collaborative practice in or near the point of delivery of care.

The type of condition (acute/chronic) that provided the focus for the interprofessional education was, of course, related to the dominant focus of the service setting (Table 16).

	Chronic	Acute	Mixed	Unclear
Community-based	25 (47%)			1 (2%)
Hospital-based	6 (11%)	18 (34%)		
Mixed			1 (2%)	
Unclear				2 (4%)

Table 16: Focus and setting of interprofessional education

Duration

Table 17 shows that most of the evaluations reported interprofessional education that extended over at least eight days of interprofessional contact for the participants. This was often spread over several months. A third of the evaluated examples were of medium duration, while just 8% were of short duration (up to one day). This distribution is likely to be different from the distribution of all interprofessional education because it is more likely that longer projects will be evaluated and the findings published. A means to check the extent of this influence on the data set analysed here would be a reasonably large random survey of practitioners' involvement in interprofessional education over, say, the previous year.

Duration	**Frequency**
Long (over 7 days)	30 (57%)
Medium (2-7 days)	17 (32%)
Short (up to 1 day)	4 (8%)
Unclear	2 (4%)

Table 17: Duration of interprofessional learning experiences

8.4 Aims, underpinning theory and type of educational intervention

Table 18 shows that most of the examples of interprofessional education included in this review had stated aims that included the improvement of care or the improvement of interprofessional collaboration (and, perhaps, through this to improve care). The interprofessional learning opportunity designed to realise these aims was usually formal (Table 10), as described in section 6.4. Note, however, that the aims of the education were not given for over a fifth of the studies. This may be due to the restrictions of space in journals and the perceived interests of each journal's audience. It may be related to the service-based nature of the most of the learning events and their evaluation. Perhaps the aim of improving care was thought too obvious or too general to be stated.

While stating clear aims in a published account may not be vital to that particular message, having clear aims for an educational development greatly enhances the probability of: matching the learning experience to desired outcomes, clearly understood roles and approaches for all those involved, and of planning an effective evaluation.

Aim	Frequency
Improving interprofessional collaboration *per se*	14 (26%)
Improving the quality of care	18 (34%)
Improved collaboration *and* improved care	10 (19%)
Not stated	11 (21%)
Total	**53**

Table 18: Stated Aims of the Interprofessional Education

It was unusual for the studies to indicate that a particular educational philosophy had underpinned the design of the interprofessional learning opportunity. Those mentioned were the learning organisation (7 studies, 13%), androgogy (3, 6%), problem based learning (1, 2%), and contact theory (1, 2%).

It is disappointing that so few authors chose to discuss the educational philosophy guiding the design of the interprofessional education. It is possible that some may have given little

thought to the values and assumptions underpinning the educational development. A lack of critical awareness of this type does not necessarily produce bad interprofessional education, but clearer thinking, greater reflection, and deeper analysis offer potential contributions to the cycle of continuous quality improvement. Reporting theory-driven and pragmatic curriculum decisions will allow the community of interprofessional educators and evaluators (and many are both) to test theories and become better informed about what is practicable and efficacious.

8.5 Learning and teaching methods

All but one (2%) of the evaluations described learning and teaching methods. Nineteen evaluations (36%) concerned guideline development or improvement. Within the remaining 34 evaluations, learning based on an exchange of information and sharing of experience between participants was the most frequently described educational method (32 studies), but this was rarely the sole educational method. Learning from peers usually emerged from small-group discussion.

Twenty-four (45%) studies described participants receiving information or practical tuition from an 'expert' (for example experienced practitioner or tutor). This often took the form of a lecture or seminar. However, due to our definition of interprofessional education, any evaluation that reported didactic teaching as the *sole* educational method was excluded from the review.

Less frequently occurring educational approaches included: -

- Problem-solving activities in interprofessional groups (11 studies: 21%).
- Practice-based interprofessional learning, for example students being allocated to clinical placements in interprofessional pairs (10 studies: 19%).
- Role-play (5 studies: 9%).

Seven evaluations (13%) described learners observing professionals at work in a variety of practice settings.

The learning and teaching methods described in the evaluations indicated that (consciously or otherwise) nearly all the interprofessional learning opportunities reflected good practice in adult learning. Generally, learners were actively engaged. Acceptance of a need to learn was usually established through a problem-focus or work-related task. Reflection on practice and values was encouraged through, for example, discussion, role-play and carefully structured observation. In part, this finding is an artefact of our inclusion criteria, which demanded some interaction between learners from different professions as part of the definition of interprofessional education.

9 Evaluation approaches and quality

It is unsurprising that the scores for quality of study and quality of information within the published account were strongly correlated (Table 1, page 19). Poorly described studies are often difficult to evaluate in terms of appropriate harmony between rationale, questions and design. They usually provide insufficient evidence to be adjudged high quality in design and conduct, because of lack of evidence on which to base this judgement. It was found that reasonably well-described studies usually provided evidence of reasonable design in relation to evaluation questions, appropriate conduct and adequate analysis. They were often weaker in considering ethics, bias and significance. Nevertheless, we feel it is important to try to tease out strengths and weaknesses in both evaluation design and the reporting of studies. Through this process evaluators may be helped to see more clearly how their practice could be developed, with the objective of producing better evaluations of interprofessional education.

Table 1 shows that no study included in this review was allocated a quality score of five for either design or reporting. Unless our quality judgements have been unreasonably harsh, this outcome is a sobering thought for the interprofessional education community (ourselves included) and provides a target for improvement.

9.1 Rationale and design

It was relatively unusual for the published accounts in the higher quality studies to indicate why an evaluation had been undertaken or to indicate its orientation (14 studies, 26%). For six evaluations (11%) the orientation was reported as criterion focused and for three studies (6%) as process focused. Three evaluations were reported as part of a quality improvement cycle and a further two studies (4%) as developmental in orientation. Similar remarks to those made above in respect of reporting educational aims and philosophy can be made of the low level of reporting for the rationale or orientation of the evaluation itself. However, as

Table 12 shows, most studies reported the evaluation design.

Naturally, given our inclusion criteria, the majority of higher quality studies have sound design features, most particularly, the collection of data at more than one point in time (e.g. before, during and after the intervention). Anderson et al (1994) comment on the value of a before-and-after design, which allows time for the implantation of practice changes (Box 2). In addition, it is clear that elaborating initial results by a comparative case study design, particularly in the context of mixed results, adds to the strength of evidence. Glanz et al. (1992) comment that their allocation of six months for the pilot evaluation was insufficient (Box 3).

Anderson et al. (1994), from the United States, report the evaluation of an in-service training programme in geriatrics for six teams that continued for 18 months following the college-based training. They examined differences between teams that had and had not made changes in organisational practice. All participants were positive about the training programme as a learning experience and about its usefulness to some aspect of their practice. At 18 months, three teams had implemented a clinical service for older adults and were satisfied with their accomplishments. The other teams had not been able to do this: they had experienced economic pressures and conflicts and were dissatisfied with their accomplishments. Analysis of reasons for the difference revealed financial and community support as key features and the authors base their recommendations for future team training activities on these findings.

Box 2: Vignette 1

Glanz et al. (1992) describe the evaluation of the implementation of cholesterol management programmes in the US, following training for physicians, cholesterol counsellors and office staff. Implementation had 14 essential components and the findings revealed diversity in the level (1-5) at which these were rated as implemented by clinic staff. There were also mixed responses about the usefulness of the training, and differences of opinion about the cholesterol management programme between the professional groups. The authors discuss the value of using a partnership evaluation model that engages staff and establishes ownership of the programme in question.

Box 3: Vignette 2

9.2 Data collection

Questionnaires were the most popular method of data collection (37 studies, 70%). These ranged from complex research instruments to 'happy sheets' gathering immediate reactions to the learning experience. Hayward *et al* (1996) highlight the use of a previously validated evaluation tool for a before and after study (Box 4).

Hayward et al. (1996) report the evaluation of a training project in a rural practicum setting, for pre-registration students from a number of health care professions. They used a pre- and post-test to assess changes in perception following the practice experience. Students completed the Interdisciplinary Education Perception Scale for which validity and reliability have been established (Leucht et al. 1990). Results were analysed by professional discipline (nursing, medicine, allied health professionals) and gender for four factors related to attitudinal responses important to collaborative practice. Results showed a significant change in perception of actual co-operation and resource sharing; with males scoring significantly lower than females.

Box 4: Vignette 3

Twenty-two evaluations (42%) analysed data from the clinical audit process. Nine studies (17%) employed documentary analysis. Interviews (11 studies, 21%), non-participant observation (8, 15%) and site visits (2, 4%) were also reported as data collection methods. The majority of studies (31, 58%) reported a single data collection method, the remainder reporting between two and five methods. Mann et al. (1996) report a programme evaluation that uses methodological triangulation (observation, interview and questionnaire). This adds to the credibility of the findings, which indicate participants learning to appreciate the role of others and to be comfortable consulting colleagues.

The sources of data included interprofessional education participants (38 studies, 72%), patient/client data (for example clinical outcomes, satisfaction) (19, 36%), audit (4, 8%), tutors (2 studies, 4%) and the reflective research journal or participant observation of the evaluator (1 study, 2%). Most studies (41, 77%) drew data from a single source category, the remainder drawing upon two or three. Eight studies (15%) included control groups, including those by Leininger & Earp (1993), (Box 5) and LaSala, Hopper, Rissmeyer & Shipe (1997), (Box 6). Just one study reported random allocation of participants between intervention and control groups, the randomized controlled trial by Solberg *et al* (1998).

The rarity of designs that incorporated a control group is not unexpected. Finding suitable controls, accounting for all the key variables associated with this group, and doing this along several points in the study, introduces a considerable level of complexity. Many service providers and universities will not have the resources to sustain evaluation of this nature. However, if the evaluation aims to establish the efficacy of an intervention it is advantageous to aim for a study design that includes a control group, or even a non-statistical comparative group.

Leininger and Earp (1993) studied the impact of smoking cessation training for clinical practice teams (clerks, nurses and physicians) in US primary care clinics, using a before and after design with controls from other practices. They compared staff attitudes and self-reported behaviour in relation to smoking cessation counselling activities. They report that a single three-hour meeting did not change attitude but did increase participation in the activities. The paper comments on the limitations of the study e.g. the use of untested attitudinal scales and their uncertainty about the generalisability of the findings. One conclusion drawn is that the training may have utility for the motivated staff.

Box 5: Vignette 4

LaSala et al. (1997) discuss the outcomes of a project to establish a community-based undergraduate course for health administration, social work and nursing students. This was evaluated by a survey tool to measure attitude changes and to track career choices. Both aspects of the tool had been used previously, but no claim is made about its validity or reliability. Results from the short-term attitude survey show an increase in the more positive attitude of the students who volunteered to take the course when compared to other [MRH1]*students. The authors recognise the bias introduced by the self-selection of students.*

Box 6: Vignette 5

9.3 Data analysis

The majority of studies (36, 68%) analysed and reported only quantitative data, but only 16 of these reported tests for statistical significance (e.g. t-test, χ^2 tests, analysis of variance). The approach to data analysis was not reported in 4 papers (8%). The unit of analysis was

normally the individual interprofessional education participant (31 studies, 59%) or the organisation (13, 25%), while eight evaluations (15%) focused upon a department. Only 17 of the 50 studies that involved data collection at more than one time point discussed loss to follow up.

The majority of studies (34, 64%) do not discuss the issue of bias. However, Glanz et al. (1992) provide an example of an evaluation design that sought to safeguard against bias (

Box 3). In this case it was through multiple measures and data sources, well-defined operational variables, and the use of qualitative and quantitative data. In addition, the authors point out the limitations of their study: mainly the lack of comparative baseline data.

The dominance of designs from the positivist paradigm and quantitative data collection tools was a disappointment. Given the nature of questions that need answering in relation to interprofessional education, we would have liked to find a greater number of studies based in the interpretive and critical paradigms, more emphasis on process, and more examples of qualitative data collection. That these did not occur may be partially due to funding and publication bias. We feel that there is much to be learnt from ethnographic, phenomenological and action research studies that focus on the social ingredients of successful interprofessional education from multiple perspectives. They will help to identify the contextual mechanisms for the impact of interprofessional education. The NHS South West Regional Office programme of evaluated three-year development projects *'Achieving health and social care improvements through interprofessional education'* (summarised in Knasel 2002) can be regarded as an example of good practice in this respect.

It is comparatively difficult to secure funding for qualitative or mixed methods studies of educational interventions for health and social care. Resistance remains from the historical dominance of modernism, the world-view that underpins traditional scientific enquiry. Indeed, a simplistic interpretation of the principles of evidence-based care could extend the breadth and depth of this resistance. Another factor militating against securing funding to research interprofessional education is the position of 'falling between stools'. Health or social care organisations may be reluctant to fund educational studies. This may be because such studies are perceived as one step removed from client care, although some of the studies reported here have demonstrated a direct impact on patent/client outcomes. Educational

purchasers may be reluctant to fund studies of post-compulsory education, or studies of education outside educational institutions, or studies of professional groups that have their own research funding streams. Preparing successful bids to conduct high quality studies of interprofessional education may also be hampered by the sparse distribution of well-qualified educational researchers among interprofessional education initiatives. There is a need to address the multi-faceted barriers to high quality research of interprofessional education.

10 Reported Outcomes

As we have previously indicated, outcomes affecting participating learners, service users and organisations were all considered to be authentic and of value. In this chapter we report on the findings of the review from the perspective of the six levels of outcomes shown in Figure 2 (page 14).

The reported outcomes from evaluations of interprofessional education in the higher quality studies were assessed to be in one of four categories: positive, mixed, neutral and negative (Table 19). Firstly, the dominance of positive findings is clear. Secondly, Table 19 shows that half the evaluations reported outcomes related to participant evaluation of the interprofessional education. The acquisition of knowledge and skills, and changes in organisational practice were the next most commonly identified outcomes. Fewer studies measure attitudinal changes, benefits to patients/clients, or changes in practitioner behaviour.

The pattern of reporting outcomes partly reflects the relative ease of gathering certain types of data and partly reflects the foci of interprofessional learning opportunities. The reported outcomes of evaluations of pre-qualification interprofessional education were concentrated at levels 1 and 2a.

	Level	Positive	Mixed	Neutral	Negative
1	reaction	27 (51%)			
2a	attitudes/perceptions	14 (27%)		2 (4%)	
2b	knowledge/skills	24 (45%)			
3	behaviour	12 (23%)		1	
4a	organisational practice	21 (40%)	3 (6%)	1 (2%)	
4b	patient benefit	9 (17%)	4 (8%)	1	

Table 19: Outcomes Reported in Evaluations

Most studies reported outcomes at more than one level (Table 20). For example, 16 of the 27 studies that reported outcomes at level 1, learners' reactions, also reported outcomes at level 2b, changes in skill or knowledge. These studies tended also to report at level 2a, changes in attitude or perception. In many cases, the same research tool (for example, a questionnaire or

semi-structured interview) can be used to collect information on these issues. Thus the co-incidence of reporting on these matters is unsurprising.

Level	1	2a	2b	3	4a	4b
1	27	12	16	9	7	4
2a		16	9	4	4	4
2b			24	5	4	2
3				13	8	2
4a					25	8
4b						14

Table 20: Co-incidence of Outcomes

There is a second cluster within, containing most of the work-based, quality improvement, interprofessional education initiatives. Most of the evaluations that reported changed behaviour (level 3) also reported changes in the organisation or delivery of care (level 4a). Likewise, most of the studies that reported benefits to patients/clients (level 4b) also reported changes in the organisation or delivery of care (level 4a). This is because changed practitioner behaviour, changed service delivery and changed outcomes for patients/clients are strongly interrelated.

In the following sections we discuss the reported outcomes at each level of our model, initially focusing on positive changes. There follows an examination of studies that report either mixed or neutral outcomes. Finally, we comment on the lack of reported negative outcomes.

10.1 Reaction to the interprofessional education

Half of the higher quality evaluations report the participants' reactions to the interprofessional education. In the other studies this information may, of course, have been obtained but omitted from the published account: perhaps to comply with publishers' word limits. We would however, urge that those involved in delivering interprofessional education

assess the learner's reaction to assist quality improvement. An unsatisfactory learning experience is unlikely to yield the desired learning and/or behavioural change. If the unsatisfactory learning experience is interprofessional in nature, the negative feelings it produces in participants may become more generally associated with interprofessional collaboration: a confusion of message and mode of transmission. Therefore, it is particularly important that interprofessional education offers a high quality learning experience.

Reaction was usually gauged from feedback questionnaires. All the reported reactions were positive and they included: -

- Rating the educational experience, see, for example, Stark et al. (1984).
- Appreciating facilitator input, see, for example, Lia-Hoagberg, Nelson & Chase. (1997).
- Enjoying the interprofessional interaction, see, for example, Greene, Cavell & Jackson (1996).

10.2 Changes in attitude and perception

Where changes in attitudes towards teamwork or other professionals had been assessed it was, with one exception, by questionnaires. For example see, Carpenter (1995), Finset *et al* (1995) and Nash & Hoy (1983). Exceptionally Long (1996) reports on the use of pre and post intervention interviews to evaluate interprofessional education. Very few studies provided information about the development, reliability or validity of their instruments. Exceptions to this were two that used previously validated questionnaires LaSala et al. (1997): Box 6, Hayward *et al* (1996): Box 4 and Carpenter (1995), where theory informed the study design.

Most studies measured changes by comparing respondents' attitudes before and after the interprofessional education, for example, Parsell et al. (1998) and Bickler (1994). In only one study was an attempt made to assess attitude change with the weaker post-intervention design (Rutter & Hagart 1990).

Four studies (Parsell et al. 1998; Finset et al. 1995; Anderson, 1994; Nash & Hoy 1983) used follow-up measures to monitor persistence of change. This is clearly important as it too easy

to assume that desired change will last. Some studies did not specify the aspects of interprofessional attitudes measured (e.g. LaSala et al. 1997: Box 6; Barber et al. 1997). Others covered a range of issues of working collaboratively, such as: -

- Confidence in teamwork (Nash & Hoy 1983).
- Views on breadth of life experience, academic quality, professional competence of other professionals (Carpenter 1995).
- Roles and function (DePoy, Wood & Miller 1997).
- Satisfaction with team function, clarify of team ideology, mutual responsibility for care (Finset et al. 1995).
- Liaison (Rutter & Hagart 1990).
- Attitudes to team importance, authority, trust, importance of the success of teams, satisfaction with team accomplishment (Anderson 1994).

10.3 Changes in knowledge and skills

Twenty- four studies (45%) reported changes in skill or knowledge relating to collaborative practice. This was usually gauged from questionnaire responses (21 studies) although some use was made of interviews, observation and documentary analysis of participants' reflective diaries or logbooks. As should be the case, nearly all the evaluations reporting outcomes at this level adopted some variation of before-and-after or longitudinal design. It would be difficult to reach a trustworthy conclusion of a change in knowledge or skill without collecting data at more than one time point. Reported changes in knowledge or skill included: -

- Improved knowledge of the nature of interprofessional teamwork (Dienst & Byl, 1981).
- Enhanced understanding of the roles and responsibilities of other health care professionals (Mann *et al*, 1996; Parsell *et al*, 1998).
- Development of teamwork skills such as interprofessional communication (Madsen *et al*, 1998; Rutter & Haggart, 1990).

10.4 Changes in behaviour

Typically, studies that included changes in behaviour (13, 25%) reported outcomes that included improved interprofessional cooperation and communication (e.g. Skovholt et al. 1994), and the development of closer links between participants (e.g. Kristjanson et al. 1997). However, we found the conceptualisation and operationalisation of this outcome variable problematic. The reporting of this outcome tended to be anecdotal rather than robustly measured. There was a particular problem with studies not establishing robust baseline or comparative data in order that changes in behaviour could be securely identified. Interpreting the findings from these studies was further complicated as it was difficult to disaggregate reports of changes in individual behaviour from changes in organisational practice.

10.5 Changes to organisational practice

Twenty-five studies (48%) reported changes in organisational practice following interprofessional education. These changes were usually detected using data from the clinical audit process or from questionnaire responses. Examples of these changes are shown in Box 7.

Overdyk et al. (1998) reported reduced 'turn-over' time for first patient of the day in the operating theatre, following discussion of delayed cases and reasons for the delay by the Operating Room Committee. Measures included average anaesthesia ready time and procedure start time, improved turn-over time and reduced empty room time

Hunter & Love (1996) showed that aggressive incidents in the dining room of a forensic psychiatric hospital were reduced as result of actions implemented by the staff team. They reviewed mealtime behaviour and conducted a patient survey. The changes were: replacing metal utensils with plastic; playing therapeutic music; permitting privileged patients to leave the dining room after eating; opening the main courtyard and gym during meals; and training food-serving staff in therapeutic communication. Outcome measures included reduced number of violent incidents, elimination of attacks with silverware and reduction of nursing time supervising mealtimes.

Bultema et al. (1996) evaluated the design and implementation of psychiatric clinical pathways for geriatric patients with depression that were based upon consultation between professions. The pathways achieved positive quality and fiscal outcomes. They used before and after measures including percentage of medical consultations, percentage of examinations by internist or practice nurse within 24 hours, staff interaction with patients, and staff contact with after-care agencies.

Hickey et al. (1996) report a set of discharge related improvements using improvement cycles designed, implemented and evaluated by a hospital team. These included; clarifying for the patient who was responsible for discharge planning decisions, preparing patients for discharge, co-ordination of dispensing of medications, and recording information regarding follow-up appointments for the benefit of carer.

Gunn et al. (1995) report on more appropriate admissions to a detoxification unit and appropriate admission to a medical ward in complicated cases resulting from the implementation of a clinical pathway by an action team.

Box 7: Examples of Changes in Organisational Practice

10.6 Benefits to patients/clients

All the evaluations reporting outcomes at this level of our model collected data at more than one time point and/or had control groups. Four broad approaches were employed, singly or in combination:

- The largest group focused on obtaining patient satisfaction; Madsen et al. (1988), Falconer et al. (1993), Bickler et al. (1994), Townes et al. (1995), Brown (2000), Hickey et al. (1996). Data collection methods included questionnaires or interviews with patients or relatives.
- Falconer et al. (1993) and Glanz et al. (1992) used pre-validated measures of mobility, communication and sphincter management.
- Other measures (e.g. forms recording students' evaluations of patient performance; reported survival rates) were employed by Dienst & Byl (1981), Falconer et al. (1993) and Brown (2000).
- Length of stay was another important marker of impact of interprofessional learning. In the context of three studies, Falconer et al. (1993), Gunn et al. (1995), Bultema et al. (1996), it implied a better organisation of care as well as better, more appropriate care.
- Only one study (Hunter & Love 1996) showed a change of patient behaviour as a result of change processes instituted after a team review.
- Glanz et al. (1992) and others used a specific clinical outcome, such as changes in serum cholesterol.

10.7 Reports of neutral or mixed outcomes

Five studies reported no change (neutral) and seven studies reported no change and positive outcomes (mixed). Box 2, Box 3 and Box 4 summarise studies where there are mixed outcomes between the participants of the interprofessional education. In each of these the differences in outcomes have different characteristics. Anderson et al. (1994) report

differences between teams at level 4a, Glanz et al. (1992) differences between the practitioners at levels 4a&b and Hayward et al. (1996) highlight gender differences at level 2a. Leininger & Earp (1993) in Box 5 is an example of outcomes that are positive at one level (in this case by changing behaviour) but neutral at another (no change in attitude). However, measuring attitudinal change is challenging and these authors point out the untested nature of their instruments. These studies also show some positive features of evaluation design and reportage (see chapter 9).

10.8 Reports of negative outcomes

We found no papers that reported wholly negative outcomes. This is perhaps unsurprising, given the tendency in the academic press to share good news and the reluctance to highlight what may be perceived as failures. This is unfortunate as much can be learnt from disappointments. For people to share these willingly there is a need to reduce the 'blame culture' that pervades education as much as practice in health and social care.

11 Conclusions

This critical review of evaluations of interprofessional education, based on a systematic review of Medline, CINAHL and the British Education Index, has revealed much about the nature, outcomes and evaluation of interprofessional education and the dissemination of findings from these evaluations.

The studies we found mostly focused on post-registration continuing professional development. The interprofessional education usually occurred within the workplace or an employer's training facilities. Fewer than 30% of studies included pre-registration students, but the location of this interprofessional education was often a service delivery setting rather than the university. The post-registration interprofessional education could be subdivided into traditional staff development based on, for example, workshops and short-courses, or on the other hand, interprofessional education that occurred as a by-product of a quality improvement initiative. Thus we see interprofessional education and collaboration promulgated both as a cause in its own right and as a problem-solving strategy.

The quality of studies varied and we found it useful to focus on the higher quality studies from chapter 7 onwards. This data set is largely North American, evenly divided between hospital and community settings and with a highly correlated division of focus between acute and chronic conditions. The learning experience was almost always formal (e.g. workshop, seminar), although not usually overtly underpinned by any particular educational theory, and generally of medium or long duration. Nursing and medicine were the most frequently represented professions, reflecting the size and role diversity of these professions relative to others in health and social care.

Within the focused set of higher quality studies the evaluation designs were dominated by variants of before-and-after studies and longitudinal studies. These do seem to be the most profitable way of investigating processes and outcomes associated with interprofessional education. However, convincingly demonstrating cause and effect remains a problem for

many studies. Few studies addressed the longevity of any changes detected. More prospective studies and more longitudinal studies are required.

Studies based on quantitative data with limited interpretation were in the majority. We now need more interpretive and critical studies. Although expensive and relatively more difficult to publish, there is much to be gained from well-conducted qualitative studies. Since most interprofessional education initiatives are multi-faceted, more mixed methods studies would also be advantageous. It is comparatively difficult to secure funding for qualitative or mixed methods studies of educational interventions for health and social care. A range of contributory factors was discussed at the end of section 9.3. There is a need to address the multi-faceted barriers to high quality research of interprofessional education.

The higher quality studies provided evidence of reasonable design in relation to evaluation questions, appropriate conduct and adequate analysis. They were often weaker in considering ethics, bias and the significance of their findings (statistical, practical, policy, etc.).

We grouped the reported outcomes of interprofessional education into six categories: learners' reactions, changes in attitude or perception, changes in knowledge or skill, behavioural changes, changes in the organisation or delivery of care, benefit to patients or clients. Most studies reported outcomes at more than one level. Studies evaluating university-based interprofessional education tended to focus on learners' reactions, attitudes and perceptions, knowledge and skill. This is not surprising. These learning experiences are about consciousness raising, preparation for future practice, and a contribution to professional socialisation: an investment in the future. On the other hand, studies evaluating interprofessional education that occurred to augment quality improvement initiatives tended to focus on behavioural change, organisational change and patient benefit. This reflects their focus on specific problem solving in practice, *here* and *now*.

The reference list for this report reveals the tremendous variety of journals publishing evaluations of interprofessional education. These studies are often located with other work relating to the client group in question, or within the quality-improvement literature.

An inevitable question set by any review of the literature is about the generalisability of the review findings. In this case we may ask how representative the studies of interprofessional

education published in peer-reviewed health- or education-related journals are likely to be of the whole picture. Two conclusions can be drawn in this respect. The duration of the published initiatives is likely to be longer than average: that may increase the probability of securing and detecting certain outcomes. Secondly, competently evaluated initiatives may be better planned and/or better funded than the average.

There is a stronger culture of evaluation of social programmes, including education, in the US than in, for example, the UK. If expectations and the allocation of funding discourage the sound evaluation of interprofessional education, educational policy makers and providers will continue to make decisions from a relatively weak evidence base. Alternatively, they will be reliant upon evidence from a context which may have a different value system and which operates in different social and political contexts. In the UK, greater investment is needed in evaluating interprofessional learning, across the spectrum of contexts described in the studies we have reviewed. Such evaluations would contribute to our knowledge about the place and role of interprofessional education in professional curricula. These evaluations will also provide valuable evidence about effective curriculum design and inform educators about how to maximise learning outcomes.

All educational innovators should operate a *plan-do-study-act* cycle to ensure high quality, well-targeted provision. However, the practice and academic communities do not need a published account of every interprofessional learning experience. Data saturation would soon be achieved. What is required is: -

- A smaller number of comprehensive evaluations of different types of interprofessional education.

- Evaluation of innovation, in the pedagogy and evaluation.

- Prospective studies with lengthy follow-up periods.

Enquiries of this nature will ensure that the interprofessional education practice of the future is informed by robust evidence for effectiveness across the wide range of provision that this review has demonstrated exists.

12 References

Anderson LA, Persky NW, Whall AL, Campbell R, Algase DL, Gillis GL, & Halter JB (1994), Interdisciplinary team training in geriatrics: reaching out to small and medium size communities, *The Gerontologist*, vol. 34, pp. 833-38.

Barber G, Borders K, Holland B, & Roberts K (1997) Life Span Forum: an interdisciplinary training experience, *Gerontology & Geriatrics Education*, 18(1): 47-59.

Barr H (2002) Interprofessional *education today, yesterday and tomorrow: a review*, London: LTSN HS&P

Barr H, Freeth D, Hammick M, Koppel I, & Reeves S (2000) Evaluations *of interprofessional education: a United Kingdom review for health and social care*, London: CAIPE/BERA

Barr H, Hammick M, Koppel I, & Reeves S (1999a) Evaluating Interprofessional Education: Two Systematic Reviews for Health and Social Care, *British Educational Research Journal,* vol. 25, no. 4, pp. 533-44.

Barr H, Hammick M, Koppel I, & Reeves S (1999b) Systematic Review of the Effectiveness of Interprofessional Education: Towards Transatlantic Collaboration, *Journal of Allied Health*, vol. 28, no. 2, pp. 104-8.

Barr H & Shaw I (1995) Shared *learning: selected examples from the literature* London: CAIPE

Barr, H & Waterton S (1996) Interprofessional *education in health and social care in the United Kingdom: report of a CAIPE survey,* London: CAIPE

Bickler, B (1994) Putting patient-focused care into practice, *AORN Journal*, vol. 60, no. 2, pp. 242-245.

Brown S (2000) Outcomes analysis of a pain management project for two rural hospitals, *Journal of Nursing Care Quality*, 14(4):28-34.

Bultema, JK, Mailliard L, Getzfrid MK, Lerner R D, & Colone M.(1996) Geriatric patients with depression Improving outcomes using a multidisciplinary clinical path model, *Journal of Nursing Administration*, vol. 26, no. 1, pp. 31-38.

Carpenter J (1995) Interprofessional education for medical and nursing students: evaluation of a programme, *Medical Education*, vol. 29, no. 4, pp. 265-272.

Chessman A, Bellack JP, Cleghorn GD Kennedy DB, Lahoz, MR, Slaughter SC, & Mahler WN (1996) Institutionalizing continuous improvement in South Carolina: taking it "bird by bird", *Joint Commission Journal on Quality Improvement*, vol. 22, no. 3, pp. 177-187.

DePoy E, Wood C, & Miller M (1997) Educating rural allied health professionals: an interdisciplinary effort, *Journal of Allied Health*, vol. 26, no. 3, pp. 127-132.

Dienst ER. & Byl N (1981) Evaluation of an educational program in health care teams, *Journal of Community Health*, vol. 6, no. 4: 282-298.

Falconer JA, Roth E J, Sutin JA, Strasser DC & Chang RW (1993) The critical path method in stroke rehabilitation: lessons from an experiment in cost containment and outcome improvement [see comments], *Quality Review Bulletin*, vol. 19, no. 1: 8-16.

Finset A, Krogstad J M, Hansen H, Berstad J, Haarberg D, Kristansen G, Saether K & Wang MD (1995) Team development and memory training in traumatic brain injury rehabilitation: two birds with one stone, *Brain Injury*, vol. 9, no. 5, pp. 495-507.

Freeth D & Nicol M (1998) Learning clinical skills: an interprofessional approach, *Nurse Education Today*, vol. 18, no. 6, pp. 455-461.

GlanzK, Brekke M, Harper D, Bache-Wiig M. & Hunninghake DB (1992) Evaluation of implementation of a cholesterol management program in physicians' offices Health *Education Research*, 7(2):151-64, 1992 June 7, 151-164.

Greene, R. J., Cavell, G. F., & Jackson, S. H. (1996), Interprofessional clinical education of medical and pharmacy students, *Medical Education*, vol. 30, no. 2, pp. 129-133.

Gunn SR, Hanisch P & Wood D (1995) CQI action team: responding to the detoxification patient, *Joint Commission Journal on Quality Improvement*, vol. 21, no. 10, pp. 531-540.

Hayward KS, Powell LT, & McRoberts J (1996) Changes in student perceptions of interdisciplinary practice in the rural setting, *Journal of Allied Health*, vol. 25, no. 4, pp. 315-327.

Headrick LA, Knapp M, Neuhauser D, Gelmon S, Norman L, Quinn & Baker R. (1996) Working from upstream to improve health care: the IHI interdisciplinary professional education collaborative, *Joint Commission Journal on Quality Improvement*, vol. 22, no. 3, pp. 149-164.

Hickey M, Kleefield S S, Pearson S D, McCabe Hassan S, Harding M, Haughie P, Lee TH. & Brennan TA (1996) Payer-Hospital Collaboration to Improve Patient Satisfaction with Hospital Discharge, *Joint Commission Journal on Quality Improvement*, vol. 22, no. 5, pp. 336-344.

Hunter M & Love C (1996) Total Quality Management and the reduction of inpatient violence and cost in a forensic psychiatric hospital, *Psychiatric Services,* vol. 47 pp. 751-54.

Kirkpatrick DL (1967) in: Craig R & Bittel L (Eds) *Training and Development Handbook*, Vol. McGraw-Hill: New York.

Knasel E / Learners First (2002) *Learning and working together: achieving health and social care improvements through interprofessional education* Bristol: NHS South West Regional Office.

Koppel I, Barr H, Reeves S, Freeth D & Hammick M (2001) Establishing a Systematic Approach to Evaluating the Effectiveness of Interprofessional Education, *Issues in Interdisciplinary Care.* vol. 3, no. 1: 41-9.

Kristjanson L, Dudgeon D, Nelson F, Henteleff P, Balneaves L (1997) Evaluation of an interdisciplinary training program in palliative care: addreassing the needs of rural and northern communities. *Journal of Palliative Care* vol. 13, no. 3, pp5-12.

LaSala KB, Hopper SK, Rissmeyer DJ & Shipe D P (1997) Rural health care & interdisciplinary education, *Nursing and Health Care Perspectives*, vol. 18, no. 6, pp. 292-298.

Leininger L & Earp JL (1993) The effect of training staff in office based smoking cessation counseling, *Patient Education & Counseling*, vol. 20, no. 1, pp. 17-25.

Leucht RM, Madson MK, Taugher MP, & Petterson J (1990) Assessing perceptions: design and validation of an interdisciplinary education perception scale, *Journal of Allied Health,* 19(2): 181-191

Lia-Hoagberg B, Nelson P & Chase RA (1997) An interdisciplinary health team training program for school staff in Minnesota, *Journal of School Health*, vol. 67, no. 3, pp. 94-97.

Long S (1996), Primary health care team workshop: team members' perspectives, *Journal of Advanced Nursing*, vol. 23, no. 5, pp. 935-941.

Madsen M K, Gresch AM, Petterson B & Taugher MP (1988) An interdisciplinary clinic for neurogenically impaired adults: a pilot project for educating students, *Journal of Allied Health*, vol. 17, no. 2, pp. 135-141.

Mann K V, Viscount P W, Cogdon A, Davidson K, Langille D B, Maccara M E (1996) Multidisciplinary learning in continuing professional education: the Heart Health Nova Scotia experience, *Journal of Continuing Education in the Health Professions* vol. 16, no. 1, pp. 50-60.

Mitchell L, Fife S, Chothia AA, Leong D, Dixon S, Airola A, Stickney J, Mueller J, Ruvalcaba R & Neuman L (1996) Three teams improving thrombolytic therapy, *Joint Commission Journal on Quality Improvement*, vol. 22, no. 6, pp. 379-390.

Nash A & Hoy A (1993) Terminal care in the community--an evaluation of residential workshops for general practitioner/district nurse teams, *Palliative Medicine*, vol. 7, no. 1, pp. 5-17.

Overdyk FJ, Harvey SC, Fishman RL & Shippey F (1998) Successful strategies for improving operating room efficiency at academic institutions, *Anesthesia & Analgesia*, vol. 86, no. 4, pp. 896-906.

Parsell G, Spalding R & Bligh J (1998) Shared goals, shared learning: evaluation of a multiprofessional course for undergraduate students, *Medical Education*, vol. 32, no. 3, pp. 304-311.

Pawson R & Tilley N (1997) *Realistic Evaluation,* Sage Publications: London.

Reeves S, Koppel I, Barr H, Freeth D, Hammick M (In press, July 2002) 12 tips for undertaking a systematic review *Medical Teacher*

Rutter DR & Hagart, J (1990) Alcohol training in south-east England: a survey and evaluation, *Alcohol & Alcoholism*, vol. 25, no. 6, pp. 699-709.

Skovholt C, Lia-Hoagberg B, Mullett S, Siiteri RK, Vanman R, Josten, McKay C & Oberg CN (1994) The Minnesota Prenatal Care Coordination Project: successes and obstacles, *Public Health Reports*, vol. 109, no. 6, pp. 774-781.

Solberg, L. I., Kottke, T. E., & Brekke, M. L. (1998), Will primary care clinics organize themselves to improve the delivery of preventive services? A randomized controlled trial, *Preventive Medicine*, vol. 27, no. 4, pp. 623-631.

Stanford & Yelloly et al (1994) *Shared learning in child protection: report of the ENB/CCETSW training and evaluation initiative to develop a model of shared learning within child protection courses* London: ENB

Stark R, Yeo G, Fordyce M, Grudzen M, Hopkins J, McGann L & Shepard (1984) An interdisciplinary teaching program in geriatrics for physician's assistants, *Journal of Allied Health*, vol. 13, no. 4, pp. 280-287.

Stein J & Brown H (1995) 'All in this together': an evaluation of joint training on the abuse of adults with learning disabilities, *Health and Social Care in the Communit,y* vol. 3, no. 4, pp 205-214

Thackwray B (1997) *Effective evaluation of training and development in higher education,* London: Kogan Page.

Townes C, Petit B & Young B (1995) Implementing total quality management in an academic surgery setting: lessons learned, *Swiss Surgery,* no. 1, pp. 15-23.

Zwarenstein M, Atkins J, Barr, Hammick M, Koppel I & Reeves S (1999) A Systematic Review of Interprofessional Education, *Journal of Interprofessional Care, v*ol. 13**,** no. 4, pp. 417-24.

Zwarenstein M, Reeves S, Barr H, Hammick M, Koppel I & Atkins J (2001) Interprofessional education: effects on professional practice and health care outcomes (Cochrane Review)

http://www.update-software.com/ccweb/cochrane/revabstr/ab002213.htm

13 Appendix I: Search strategies

Medline search strategy

1 INTER-PROFESSION* or INTERPROFESSION*
2 INTER-DISCIPLIN* or INTERDISCIPLIN*
3 INTER-OCCUPATION* or INTEROCCUPATION*
4 INTER-INSTITUTION* or INTERINSTITUTION*
5 INTER-AGEN* or INTERAGEN*
6 INTER-SECTOR* or INTERSECTOR*.
7 INTER-DEPARTMENT* or INTERDEPARTMENT*
8 INTER-ORGANISATION* or INTERORGANISATION*
9 INTERPROFESSIONAL RELATIONS
10 TEAM*
11 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10
12 MULTI-PROFESSION* or MULTIPROFESSION*
13 MULTI-DISCIPLIN* or MULTIDISCIPLIN*
14 MULTI-INSTITUTION* or MULTIINSTITUTION*
15 MULTI-OCCUPATION* or MULTIOCCUPATION*
16 MULTI-AGEN* or MULTIAGEN*
17 MULTI-SECTOR* or MULTISECTOR*
18 MULTI-ORGANISATION* or MULTIORGANISATION*
19 PROFESSIONAL-PATIENT RELATION*
20 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19
21 11 or 20
22 EDUCATION* or TRAIN* or LEARN* or TEACH* or COURSE*
23 QUALITY ASSURANCE or TQM or CQI or GUIDELINE DEVELOPMENT
24 22 or 23
25 20 and 24
26 STUDENT PERFORMANCE APPRAISAL
27 COURSE EVALUATION
28 PROGRAM* EVALUATION
29 EVALUATION RESEARCH
30 EVALUATION METHODS
31 HEALTH CARE OUTCOME*
32 SOCIAL CARE OUTCOME*
33 EDUCATION* OUTCOME*
34 LEARNING OUTCOME*
35 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34
36 25 and 35

The Medline search strategy was adapted for the CINAHL and BEI searches to reflect differences in the indexing of terms between databases.

14 Appendix II: Data Abstraction Sheets

Quantitative Data Abstraction Sheet – version 6[e]

CRITERIA	COMMENTS
Ref. No:	
Citation	
Type (including grey literature)	
Educational Initiative	
Aim/objective of IPE (Implicit/explicit)	
Type of IPE	
Content	
Duration	
Method of learning/ teaching	
Location	
Participants (no. & type)	
Sector	
Level / stage	
Qualification	
Context	
Rationale for IPE (implicit or explicit)	
Outcomes	
Explicit/implicit	
Level 1	
Level 2a	
Level 2b	
Level 3	
Level 4a	
Level 4b	
Other/unspecified	
Methods of Evaluation	
Aim of Evaluation (Implicit/explicit)	
Research Design	
Data collection method	
Source of data	
Data analysis method	
Number of groups (in study)	
Unit of study (1,2, or more individuals)	
Method of allocation	
Allocation concealment	
Blinding	
Power calculation	
(Original) Sample size	
Loss to follow up	
Significance measures	
Reported biases	
Strength of design	
Strength of no.	
Quality of study	
Quality of info	
Overall weighting	

[e] NB Each data abstraction sheet has an associated set of guidance notes for its completion. These notes remind team members of negotiated decisions and agreed criteria.

Qualitative Data Abstraction Sheet – version 2

CRITERIA	COMMENTS
Ref. No:	
Citation	
Type (including, grey literature)	
Educational Initiative	
Aim/objective of IPE	
Type of IPE	
Content	
Duration	
Method of learning/ teaching	
Location	
Participants (number & type)	
Sector	
Level / stage	
Qualification	
IPE Context	
Rationale for IPE	
Outcomes	
Explicit/implicit	
Level 1: Reaction	
Level 2a: Attitudes	
Level 2b: Skills	
Level 3: Behaviour	
Level 4a: Practice	
Level 4b: Patients	
Other/unspecified	
Methods of evaluation	
Aim of Evaluation	
Sampling	
Data collection	
Data analysis	
Research relations	
Ethics	
Findings	
Transferability	
Relevance & Usefulness	
Quality of study	
Quality of information	
Overall weighting	

The Higher Education Academy Health Sciences and Practice

King's College London
Room 3.12
Waterloo Bridge Wing
Franklin-Wilkins Building
150 Stamford Street
London SE1 9NH

Tel:020 7848 3141
Fax: 020 7848 3130
http://www.health.heacademy.ac.uk