The Effectiveness of Youth Mentoring Programmes in New Zealand

July 2010

Reviewers:

Susan Farruggia PhD

Pat Bullen PhD

Ann Dunphy MA

Frank Solomon MA

Efeso Collins MA

This systematic review is part of the 'What Works in Youth Development' series, and was co-funded by the Health Research Council of New Zealand and the Ministry of Youth Development as part of the Health Research Council of New Zealand's Partnership Programme.

Reports from Auckland UniServices Limited should only be used for the purposes for which they were commissioned. If it is proposed to use a report prepared by Auckland UniServices Limited for a different purpose or in a different context from that intended at the time of commissioning the work, then UniServices should be consulted to verify whether the report is being correctly interpreted. In particular it is requested that, where quoted, conclusions given in UniServices reports should be stated in full.

Contact information for principal investigator:
Dr Susan P. Farruggia
School of Teaching, Learning and Development
Faculty of Education
The University of Auckland
Private Bag 92601
Symonds Street
Auckland 1150 New Zealand
Telephone: + 09 623 8899 ext. 48326
E-mail: s.farruggia@auckland.ac.nz

Contributors to the review:
Susan Farruggia (SF), with the assistance of Pat Bullen (PB), wrote the protocol. Research assistant Joy Davidson (JD) conducted the searches, under the guidance of SF; Bev Cassidy-MacKenzie, under the guidance of Ann Dunphy (AD) and SF, contacted the agencies. With SF taking the lead, all named Reviewers, SF, PB, AD, Frank Solomon, (FS), and Efeso Collins (EC), and JD wrote the text of the review.

Acknowledgements:
The review was conducted with support from the New Zealand Ministry of Youth Development and the Health Research Council of New Zealand Partnership Programme.

Potential conflicts of interest:
Efeso Collins is a Programme Director of Project K, a personal friend of the Founder and General Manager of Affirming Works, and he previously held a position on the MATES Advisory Board. Ann Dunphy is a current member of the MATES Advisory Board.

Table of Contents

SYNOPSIS **4**

ABSTRACT **5**

Background 5
Objectives 5
Search strategy 5
Selection criteria 5
Data collection and analysis 5
Main results 6
Reviewers' conclusions 6

BACKGROUND FOR THE REVIEW **7**

Vulnerable youth in New Zealand 7
Mentoring as a social intervention 8
Mentoring in New Zealand 9
The current study 11
Results 16

DISCUSSION **22**

CONCLUSIONS AND RECOMMENDATIONS **26**

ACKNOWLEDGEMENTS **27**

POTENTIAL CONFLICTS OF INTEREST **27**

REFERENCES **28**

TABLES **49**

APPENDICES **55**

Appendix A: Coding sheets 55
Appendix B: Review Protocol 664

Synopsis

There are many different interventions currently available that can be used as a preventative measure for at-risk youth.

One common intervention involves partnering youth with a more mature role model in a mentoring relationship.

While research in youth mentoring is extensive overseas, little research has been done into its effectiveness in New Zealand, despite its growth in the past 20 years. This review is based on evaluations of mentoring programmes in New Zealand.

Of the 74 potential studies identified in searches, 26 met the inclusion criteria for this review. Results showed there are currently 23 active mentoring programmes operating in New Zealand; however, only 35% have conducted evaluations examining the effectiveness for mentees.

Overall, 88% of the programmes included in this review showed some level of effectiveness, although the results are tentative due to the varied quality of the research. Further, programmes that focused on psychological and interpersonal goals were more effective than programmes focused on educational, behavioural, vocational or cultural goals.

Programme characteristics that appeared to moderate effectiveness included: dissemination, age of programme, history of evaluation, utilising principles of best practice, component programme, type of mentoring relationship, use of peers as mentors, level of structure, expected length of mentor-mentee relationship, SES of youth, and researcher-practitioner relationship.

One important caveat to these findings is that the quality of the research was extremely variable, with a significant proportion being of poor quality.

Finally, while almost all programmes and the research conducted on those programmes were culturally appropriate to the overall New Zealand cultural context, they were, as a whole, less culturally appropriate for programmes working with Māori and Pasifika youth.

Abstract

Background

It is widely accepted that there is a section of vulnerable youth in New Zealand in need of some level of intervention. Internationally, mentoring has a long history as a social intervention, and has been growing steadily since the 1990's in New Zealand. It provides a means of social support that can be accessible to vulnerable youth. Mentoring involves the partnering of a more experienced mentor with a younger less experienced mentee in a relationship characterised by mutual trust and the growth and development of the mentee. Prior research from overseas has demonstrated that this relationship has been linked to increased well-being for at-risk youth. However, a thorough assessment of mentoring in New Zealand has yet to be conducted and is vital in order to inform policy and practice and to develop the most effective mentoring programmes for New Zealand youth.

Objectives

This systematic review has the following objectives:

1) To examine the effectiveness of youth mentoring programmes in New Zealand;
2) To identify the characteristics of successful programmes;
3) To assess the quality of the research on youth mentoring; and
4) To identify gaps in the literature and recommend directions for future research.

Search strategy

A range of educational, psychological, social science and New Zealand databases were searched with a list of search terms created with the help of a Subject Librarian. Government Ministry and Research sites, as well as Google and Google Scholar, were also searched. In addition, all mentoring programmes on the database of the Youth Mentoring Network were contacted and asked to provide evaluation reports and references for such reports. Finally, a search was conducted on all reference lists of obtained articles.

Selection criteria

Only studies that met all the below inclusion criteria were included in the review.

1) Studies were required to examine the effectiveness of the programme and to address programme outcomes.
2) Participants in the studies were required to be over the age of 6 years and under the age of 24 years, with the mean age for the study not being over 19 years.
3) Studies were to be set in New Zealand.
4) Studies were required to involve a formal mentoring programme and not to focus on informal or natural mentoring.
5) For qualitative studies to be included they needed an indicator of effectiveness reflecting change; post-test only was acceptable if change was discussed.
6) For quantitative studies to be included there needed to be an indicator of effectiveness including an indication of change or difference (e.g., pre-test post-test change or the use of a comparison group; post-test only with an indicator of effect).

Data collection and analysis

All relevant full-text studies identified during the literature search were double-coded for inclusion against the inclusion criteria. All included studies were then double-coded for bias and double-coded for data extraction. Code sheets were developed using guidelines from Littell, Corcoran and Pillai (2008) and using DuBios, Holloway, Valentine and Cooper (2002) and Tolan, Henry, Schoeny and Bass (2008) as starting points. The data extraction coding sheet covered aspects of the evaluative research design, programme characteristics and programme outcomes of each evaluation. Quantitative and qualitative studies were assessed for effectiveness independently and then merged.

Main results

Of the 74 studies originally examined for inclusion, 26 met the criteria and were retained. Results showed that while there are currently 23 active mentoring programmes in New Zealand, only 35% have conducted evaluations examining the effectiveness for mentees. Mentoring tended to be one-to-one (73%), with most mentors being an adult. Most programmes screened and trained their mentors and for the most part mentors were volunteers. Programmes typically had little or no family involvement (71%). Although most programmes targeted high or at-risk youth (76%), 18% targeted typical/low risk youth, 18% targeted mixed-risk youth. Almost all programmes identified at least one programme goal (96%), with almost all including an educational goal (96%), followed by psychological (52%) and interpersonal (52%) goals.

Overall, 88% of the programmes included in this review showed some level of effectiveness; however, this finding is tentative due to the varied quality of the research. Programmes that focused on psychological and interpersonal goals were more effective than programmes focused on educational, behavioural, vocational or cultural goals. Effective programmes typically: were more established; had a history of evaluation; utilised principles of best practice; had mentoring as a component of other interventions; had adult mentors; utilised one-to-one or mixed mentoring; were more structured; had greater expectations on the length of the mentor-mentee relationship; worked with low and mixed SES youth; and differentiated researchers from practitioners.

In terms of research methodology, 31% of evaluations were purely quantitative, 31% were purely qualitative, and 38% employed a mixed method design. One important caveat to the findings regarding effectiveness is that the quality of the research was extremely variable, with a significant proportion being of poor quality, such as having high levels of bias, not using a control group, or not utilising a pre-test post-test design. Further, many of the studies included in this review did not provide a thorough description of how the programme was delivered, consequently resulting in a large amount of missing data.

Overall, there were 14 studies that included Māori mentees, and 6 that included Pasifika mentees. Of the 14 programmes that included Māori mentees, a large proportion (50%) was rated as having completely ignored Māori culture. Of the studies that included Pasifika mentees, most fell within the moderate range of cultural appropriateness, with one rated as having completely ignored Pasifika culture.

Reviewers' conclusions

A high proportion of mentoring programmes had educational goals, yet effectiveness within this domain was varied and less so than programmes with psychological and interpersonal goals. Ineffective programmes tended to be less specialised possibly reflecting the need to become more structured with fewer programme goals. As many of the characteristics of effective programmes identified in this review map onto international principles of best practice, programmes should ensure these are incorporated within their delivery. This review highlighted that only a small proportion of active mentoring programmes have engaged in programme evaluation to assess effectiveness for youth. Further the research that has been conducted to date is clearly of varying quality, as seen by a lack of control groups, lack of a pre-test post-test design, high levels of bias, and high levels of internal evaluation. Finally, a cultural framework needs to be considered when delivering programmes to and conducting research with Māori and Pasifika youth.

Background for the Review

Vulnerable youth in New Zealand

Addressing the needs of vulnerable youth is always a high priority of any government. Government statistics in New Zealand indicate that there is a group of vulnerable youth in New Zealand society requiring preventative interventions. Areas of vulnerability for these youth are seen in educational, health and social domains. These issues tend to be particularly pertinent for Māori (indigenous people of New Zealand) and Pasifika (Pacific Island) youth, who are a special target of many interventions aimed at at-risk youth, for youth in low socio-economic areas, as well as for youth living in urban areas (75% of the total youth population, MYD, 2003).

Educationally, around 17% of all youth leave school with no qualification. This pattern is about twice as evident for Māori and Pasifika youth as it is for Pakeha (New Zealand European) youth (MYD, 2003). Government statistics also show that Māori and Pasifika youth tend to underperform academically compared to their Pakeha and Asian background peers (MOMA, 2000; MOPIA, 2003). Health-wise, there is a relatively high prevalence of mental illness among New Zealand youth, with suicide being the second most common cause of death for this sector of the population (MYD, 2003).

In terms of family, studies have found that almost half of all New Zealand children experience the separation or divorce of their parents, with just over 25% of all children and youth living in single parent families (MYD, 2003). In terms of delinquency and problem behaviour, the Youth 2007 study found that 15% of male students and 9% of female students reported having been in trouble with the police in the previous 12 months (Clark et al., 2009). Further, in the previous year 8% of students reported they had stolen something worth more than $50, 10% of students reported they had tagged or painted graffiti on someone else's property, 20% of students reported that they had deliberately damaged property that was not their own, and 40% of male students and 27% of female students reported that they had hit or physically harmed another person (Clark, et al., 2009).

While it is important to acknowledge the challenges some youth face, it is equally important to point out that every young person has potential. That is, with the right tools and nurturing environments, all youth have the ability to direct their lives in a positive way, facilitating what is referred to as positive youth development (PYD; Farruggia & Bullen, in press; Larson, 2000). A key concept of PYD is that positive change is possible through positive intervention (Lerner & Castellino, 2002), or the promotion of developmental assets (Search Institute, 2003), such as caring families and communities that provide social support.

Much research states that social support is a key protective factor for vulnerable youth (e.g., Werner & Smith, 1982). While most young people continue to find the encouraging relationships they need without the assistance of formal programmes, it can be argued that there is a significant and vulnerable group who are missing out on these crucial connections or that additional help is needed (Farruggia, Bullen & Davidson, under review). Supporting young people to develop the relationships, attitudes, values, interests and behaviours for the complexities of life cannot be done prescriptively or in the short-term. A genuine caring relationship is necessary, in which trust is built over time and the young person's strengths are encouraged.

Since the 1990's a national response to changing social circumstances has developed, due to New Zealand's realisation that a new form of structured support was needed to help young people navigate the increasing complexities of life at the end of the 20th century. One important way this social support was delivered followed the movement in the United States to set up formal mentoring programmes as a means of providing support for vulnerable youth.

Mentoring as a social intervention

The concept of mentoring has been around for centuries, as a long accepted practice used with various populations and across many different situations. However, despite the extent of its use as a social intervention, there are still many varying definitions of mentoring and much debate over what should and should not define the mentoring relationship (Allen, Rhodes & Eby, 2007).

Allen, Rhodes and Eby (2007) offered several attributes of mentoring that provide a helpful framework for understanding mentoring in spite of its varying definitions. Firstly, mentoring involves a distinctive relationship between individuals, in which each individual mentoring relationship will be unique from other mentoring relationships. In this sense, each mentorship may be unique and outcomes per mentorship may vary. Added to this, DuBois and Karcher (2005) state that this unique relationship is characterised by an emotional bond between the mentor and mentee/s which promotes an environment of trust between the individuals. Secondly, mentoring involves some gaining of knowledge by those involved in the relationship. Despite differing goals of mentoring programmes, and the various ways in which mentoring is conducted, there is always some acquisition of new information and/or skills that takes place in the mentorship. Thirdly, mentoring involves some sort of support to be given to the mentee; this support can vary from emotional and psychological to vocational and academic. This implies that the mentor has some greater experience or expertise which they are able to share with the mentee (DuBois & Karcher, 2005). Fourthly, the main goal of the mentorship is the growth and development, in whatever area, of the mentee. In this sense, the mentor guides the mentee in some area/s pertinent to the mentee. Interestingly, researchers have noted that there is often some benefit for the mentor during this relationship as well (e.g., Bullen, Farruggia, Rozas Gómez, Hebaishi & Mahmood, in press). Finally, mentoring relationships can be described as being dynamic and ever changing, as well as increasing in impact over time. For the purposes of this study, however, mentoring will be described as “an enduring relationship between a novice and an older, more experienced individual who provides guidance in a particular domain” (Evans & Ave, 2000, p. 41) which acknowledges the above attributes.

Five types of mentoring have been traditionally identified including: the traditional one-to-one mentoring pair, group mentoring, team mentoring, peer mentoring and e-mentoring (MENTOR/ National Mentoring Partnership, 2005). Traditional one-to-one mentoring involves one adult matched with one child in which the pair typically meets for at least four hours per month for at least a year. Some exceptions, such as school-based mentoring programmes, exist in which the duration is often shorter. In group mentoring, one adult mentor is joined with up to four youth to develop mentoring relationships with the mentor typically serving as a leader for group-based activities. For team mentoring, several adults work with several groups of young people. There is often fluidity in the mentoring units; however, the adult-youth ratio is typically no more than one to four. Peer mentoring occurs when a young person mentors a younger person. Typically, peer mentoring is school-based with an older student mentoring a younger student during school hours. Finally, e-mentoring involves one adult forging a relationship over the internet with one youth; sometimes they have a few face-to-face meetings. Mostly, this relationship is focused around school or career support and sometimes can serve as a bridge during the summer holiday period for other types of mentoring.

One criticism of this classification of mentoring types is that peer mentoring focuses only on the age of the mentor whereas the other four types focus on the nature of the relationship. In addition, some programmes have elements of more than one type of mentoring, such as having a one-to-one pairing in the context of the group. However, regardless of definition and type of mentoring, this relationship has been found to have significant positive effects, especially in the case of at-risk youth (DuBois et al, 2002).

Researchers have conducted studies to look at the impact of mentoring on at-risk populations. Involvement in youth mentoring programmes has been found to be associated with less absence from school, more positive attitudes toward school, greater well-being, a more positive reaction to situations

involving drugs (LoSciuto, Rajala, Townsend, & Taylor, 1996), less likelihood to start using illegal drugs and alcohol, less engagement in aggressive behaviour, decreases in wagging school, and lying to parents (Grossman & Tierney, 1998).

Evans and Ave (2000) have summarised the youth mentoring literature and have identified five mechanisms through which mentors can promote change within a young person. These include: serving as a role model to demonstrate qualities and behaviours for the young person to imitate and internalize; acting as a substitute parent for youth who either don't have a parent or whose relationship with their parents is poor; providing social support, especially emotional support; developing specific, positive skills, such as those needed in the career and work domain; and, finally, modifying undesirable behaviours, such as improving academic achievement and motivation and decreasing involvement in problem behaviour. It is important to note that not all mentors act in all of these ways; instead, mentors utilise the mechanisms that are appropriate to their mentees and are within the scope of their programme.

However, even though there is strong evidence for the effect of mentoring, much research is being done to investigate what particular parts of mentoring contribute to these improvements. DuBois and colleagues (2002) conducted a meta-analysis examining the effectiveness of one-to-one mentoring programmes in the United States. Reviewing 55 evaluations of youth mentoring programmes, they identified characteristics of the most effective programmes, or principles of best practice, including: strong relations between youth and mentors, using mentors from "helping" backgrounds, providing ongoing training and support to the mentors, involving parents, programmes that are based on both theory and research, and targeting at-risk (versus typical) youth. They found that matching on gender, race or interest, as are commonly practiced among many mentoring programmes, did not impact the effectiveness of the programme.

Much of the research on mentoring has taken place in the United States, where many wide-reaching mentoring programmes have been implemented, such as the Big Brothers, Big Sisters programme. However, mentoring is rapidly becoming popular in many countries around the world.

Mentoring in New Zealand

There is general agreement that youth mentoring was well-established in New Zealand before the term came into use. Formal youth mentoring programmes could be said to have started in New Zealand in the 1980s, when the peer support model was imported from Australia and adopted by almost all secondary schools. However, traditional apprenticeship system which came to a close around that time, had for generations supported the move into adulthood, for young men in particular. Furthermore, Māori traditions of Tuakana/Teina, where older whanau members supported younger members, pre-date European contact.

The term "mentor", first came into New Zealand usage from the business sector. Mentoring was not applied to the youth sector until the early 1990s, with the first formal youth mentoring programmes occurring in the South Island. A spontaneous national response to changing social circumstances was developing through the 1990s to provide some form of structured social support for young people. The process began with atomised development of formal programmes, while informal mentoring continued to take place in a range of communities, following long-term patterns.

Dunedin-based programme pioneers Presbyterian Support Otago, first gained a Churchill Fellowship for Jill McDonald, the then Manager of their Buddy Programme, to study mentoring programmes in the United States and United Kingdom. Upon her return to New Zealand, she, with the support of a committee of interested others, organised and hosted New Zealand's inaugural Youth Mentoring Conference in 2000. In the concluding hours of that conference, a steering group was elected to

develop the concept of a national mentoring organisation to be known as the Youth Mentoring Association of Aotearoa New Zealand (YMAANZ).

In 2001, the YMAANZ steering group organised a second national conference that took place in Blenheim. The Association was ratified by the attendees and a wider group of programme providers was elected to the Committee. Over subsequent years, YMAANZ developed some well-formulated guidelines and agreements. However, in the long term, the Committee was depleted by the resignation of significant foundation members and the Association lost impetus in the face of inevitable challenges of distance, funding, and the conflicting time and energy demands of meeting the needs of their respective programmes.

Elsewhere, initially unaware of the YMAANZ initiative, an Auckland group of interested non-programme providers had formed the Auckland Youth Mentoring Association (AYMA). This group had the advantages of the country's major population base, and as non-providers were free to focus on more modest regional needs, such as running seminars. In 2002, thanks to a partnership with the City of Manukau Education Trust, AYMA produced New Zealand's first national mentoring guide: *Youth Mentoring: An Advice Manual for Manukau and Beyond*.

Meanwhile, the development of the World Wide Web made access to international knowledge more accessible, and programme providers continued to promote youth mentoring in New Zealand. Two clear strands of specialist youth mentoring programmes thus became well known: locally-developed models such as the Buddy programmes and Project K, and New Zealand versions of notable international programmes such as Big Brothers Big Sisters. There was considerable growth in programmes of both types, thanks largely to philanthropic funding.

In 2005, the AYMA was re-established as a trust, now know as the Youth Mentoring Network (YMN), launching their National website in 2006, and then holding the first North Island-based conference on youth mentoring in 2007. Ministry of Youth Development funding then made possible the development of the *Guide to Effective Practice in Youth Mentoring New Zealand* (GYM; Youth Mentoring Network, 2009). This document provided clear links with youth mentoring in New Zealand, the Youth Development Strategy Aotearoa (Ministry of Youth Development, 2002), and international practice models. In August 2009, a further national conference was held in Auckland, building on the linkages and knowledge that had been clarified through the major consultation exercise that preceded the GYM publication. As a result of these conferences, collaboration with the Australian Youth Mentoring Network has strengthened, to include sharing of resources and alignment of respective national conferences in alternate years.

In examining the New Zealand context of mentoring, Evans and Ave (2000) suggest that practices of mentoring in the United States, where formal mentoring first originated over 100 years ago and much of the research on youth mentoring has been conducted, do not necessarily fit with the New Zealand familial/social structure. While the programmes based in the United States typically involve one-on-one relationships, this may not be appropriate for youth in New Zealand where this practice may conflict with social and cultural structures, such as whānau. Mentoring for young people in New Zealand needs to account for the cultural needs and practices of its youth. Authors of this review also note that the American emphasis on mentoring programmes for at-risk youth may be too narrow within the New Zealand context as many programmes have been established for low-risk youth.

In relation to Māori, acknowledging the importance of personal Māori identity is an issue that has to be addressed, particularly with youth (Borell, 2005). Issues around schooling and home background, parent/guardian backgrounds, and involvement in Te Reo Māori me ona Tikanga are but a few of the many variables that should be considered; however, there has been some difficulty in qualifying and quantifying the "Māoriness" of a young person and what exactly is meant by "Māori identity". According to O'Reagan (1987), Māori identity is structured around three key areas of knowledge.

These include knowing who you are, who you are related to, and your descent. By helping the young person to know of his/her whakapapa, you are providing the basis of addressing self-confidence and awareness; if this can be expressed in Te Reo Māori, the young person's Māori identity can be further strengthened. These ideas are also expressed in the work of Pere (1984) who noted that Māori identity is best understood from a holistic as opposed to an individual perspective. In illustrating this, Pere used the Te Wheke (Octopus) model showing the inter-connectedness of each tentacle/component to the other and how these sustain the whole. Further, Durie's (1998) Te Whare Tapa Wha model of health clearly identifies the interconnectedness between the four taha – Taha Wairua (spiritual side), Taha Hinengaro (cognitive side), Taha Tinana (physical side), Taha Whānau (kinship ties). An understanding of this holistic approach is vital when wanting to work successfully with Māori youth.

In addition, an understanding of how Māori hui (meetings) are run is important. Specifically, providing opportunity to self-identify within the Māori cultural context at the outset of a hui, family involvement, opening with karakia, and creating situations of involvement, rather than just being spoken to, are essential. In working with Māori youth, Hawk, Cowley, Hill and Sutherland (2001, as cited in Hammond, 2007, p. 10) suggested that effective relationships between students and teachers/mentors have three characteristics. These are "empathy and an understanding of Māori culture", "caring about the student" and "respect for the student".

Given the above, some important questions should be raised regarding mentoring of Māori youth in New Zealand. First, to what extent have mentoring programmes that include Māori youth addressed what it means to be "Māori" for both mentors and mentees? More specifically, do these programmes include among their goals strengthened cultural identity and should they? Second, to what extent does mentor training incorporate an understanding of Tikanga Māori? Third, to what extent do programmes engage whānau?

In relation to Pasifika youth, like Māori, they have a need to understand their personal/familial history. As a migrant community, the young people find themselves balancing the values of Western society with traditional values which are being reinterpreted by their parents (Anae, 2001). There are also language challenges in that many of the young people do not speak their "mother tongue" at home (Adolescent Health Research Group, 2008), as a result of a move in the 1980's by many schools to discourage Pasifika parents from speaking their language at home. Pasifika youths' needs are located in an identity/culture perspective, but also in the migrant nature of their traverse to New Zealand (Anae, 2001). It should be noted that the Pasifika community is very young, with more than a third of it being under the age of 15 years. Most Pasifika people live in Auckland with the largest population residing in Manukau City (Statistics New Zealand, 2006). There is a high proportion of Pasifika young people attending low decile schools and living in some of the poorest communities in Auckland. To that end, their identity/culture is not only affected by being a migrant community, but also by the sociological factors of having poor health outcomes and low educational achievement in school (MOPIA, 2003). These issues potentially have important implications for providers and will be included in the coding scheme for this review, such as type of delivery (traditional versus group) and the incorporation of cultural practices. It is possible that mentoring will play a pivotal role in the ability of Pasifika young people to both achieve at school and participate in NZ society, as it can provide another significant adult who will support them in seeing their aspirations come to fruition.

The current study

Given the seeming success of mentoring, particularly for at-risk youth, and its apparent establishment as a social intervention in New Zealand history, it is important that mentoring be systematically evaluated. Yet the wide use of mentoring in New Zealand has yet to be matched with this sort of thorough evaluation. Therefore, there is a current national priority and policy interest, in the area of mentoring, to conduct formal research seeking to establish effectiveness of mentoring programmes within New Zealand. The protocol for this review can be found in Appendix B.

Objectives of the Review

This systematic review has the following objectives:

1) To examine the effectiveness of youth mentoring programmes in New Zealand;
2) To identify the characteristics of successful programmes;
3) To assess the quality of the research on youth mentoring; and
4) To identify gaps in the literature and recommend directions for future research.

Methods

Criteria for considering studies for this review

Only studies that met all the inclusion criteria were included in the review. The criteria included:

Outcomes measured/programme effects

Studies were required to examine the effectiveness of the programme and to address outcomes in at least one of the following areas: emotional/psychological, problem/high-risk behaviour, academic/educational, career/employment, and social competence.

Types of participants

The review was limited to studies whose participants were over the age of 6 years and under the age of 24 years, with the mean age for the study not being over 19 years. The review was also limited to studies set in New Zealand.

Programme/intervention type

Studies were required to involve a formal mentoring programme; this could include one-to-one, group, team, peer or e-mentoring, but must not focus on informal or natural mentoring.

Research design quality

Due to the limited literature on the topic, studies with less rigorous methodologies were included, but bias was identified. In addition, both qualitative and quantitative studies were included, as long as they met the following criteria. For qualitative studies to be included they needed an indicator of effectiveness reflecting change; post-test only was acceptable if change was discussed. For quantitative studies to be included there needed to be an indicator of effectiveness including an indication of change or difference (e.g., pre-test/post-test change or the use of a comparison group; post-test only with an indicator of effect).

Search strategy

Search strategy for identifying relevant studies

The search strategy for relevant literature was conducted in four main ways. Firstly, a contact at the YMN approached all youth mentoring organisations that were part of the Network to request copies of any evaluation reports on their particular programme. It is believed that all established, active mentoring programmes in New Zealand are on the database. In-person requests of reports were made at the YMN 2009 Conference; followed by an email request of any available reports. Once reports were received, for programmes that were under the umbrella of the YMN, additional emails were sent and websites were searched to extract any information not found in the reports relating to programme characteristics.

Secondly, an extensive database search was conducted. Before this search began, a Subject Librarian was contacted and met with one of the research assistants to establish effective search strategies. The Librarian advised that due to the limited literature that would be available on the topic, it would be best to use fewer but very broad search terms in order to cover all possible reports. The Librarian also advised on what databases to search and on how to import all literature into EndNote X1 to check for duplicates. Thirdly, an internet search was conducted which covered National research sites, Ministerial websites, Google and Google Scholar. Lastly, reference lists of retained reports were checked for further reports that had not been identified by the above methods.

Search terms and databases searched
Search terms were initially based on those used by prior literature reviews and meta-analyses (e.g., Tolan et al., 2008). Consultation with a Librarian with expertise in this area resulted in a refining of this list resulting in fewer, but broader terms. The list of search terms developed with the assistance of the Librarian included: mentor*, role model; youth, young*, child*, teen*, adolescen*, juvenile; program*, evaluat*, intervent*. 'Zealand' was added as a term to all searches.

A list of databases was identified, again with the assistance of the Librarian, which would be relevant to the topic and cover a range of academic subject areas. Databases searched included: Education databases (i.e., ERIC, A Plus Education, Education Sage, Professional Development Collection, and Proquest Education Journals); Psychological databases (i.e., Psych Info, MEDLINE, Psychological and Behaviour Sciences Collection, Web of Science, and Science Direct); Social Science databases (i.e., FAMILY, Proquest Social Science Journals, Social Services Abstracts, and SAGE Sociology); New Zealand databases (i.e., Index New Zealand and Te Puna); other databases [i.e., Proquest Dissertations and Theses, The Cochrane Library, and Google Scholar (only first 100 hits on Google Scholar were examined)]. All articles found in a search were imported into EndNote X1 where duplicates were identified. Studies were not excluded based on language.

Data collection and analysis methods
Selection of studies
A total of 13,292 studies (unduplicated citations) were identified during the search. Of the studies identified during the search, two were unobtainable and a further two on-going studies were not included in this review. A total of 74 were deemed to be relevant to the review based on the citation and abstract. All relevant full-text reports that were retained during the literature search were coded using an inclusion code sheet adapted from Littell et al. (2008) and Tolan et al. (2008). This coding sheet, found in Appendix A, covered all the inclusion criteria, as listed previously, which were coded as either met or unmet for each study. This inclusion coding was done by two independent coders, a trained research assistant (RA) and the second investigator (SI), and the inter-rater agreement was 83%. All discrepancies were discussed. If agreement could not be reached, the principal investigator (PI) made the final decision. A total of 26 studies were coded as meeting the inclusion criteria for the review and are included in this study. A list of excluded studies and the reason for exclusion can be found in Table 1. A flow chart showing the selection process can be found in Figure 1.

Data Extraction and management
A data extraction coding sheet was developed for the purposes of extracting relevant information for the review from the included studies. Littell, Corcoran and Pillai's (2008) guide to systematic reviews was followed in the creation of the data extraction coding sheet. Previous literature reviews and meta-analyses were also consulted during the development of the coding sheet (i.e., Tolan et al., 2008; DuBois et al., 2002). These were adapted to fit the New Zealand context. The data extraction coding sheet covered aspects of report/research characteristics and methodologies, programme features, youth and mentee characteristics, outcome goals and measures (i.e., educational, psychological, behavioural, interpersonal, vocational, and cultural adjustment) adverse effects, timing of intervention, and quantitative and qualitative outcomes. Additionally, a description of the programme goals and corresponding indicators used in this review can be found in Table 2. An important difference in our coding scheme, in comparison to Tolan et al. (2008) and DuBois et al. (2002), was our classification of mentoring type. We opted to categorise programmes as being one-to-one, group, or mixed as it seemed to better fit the New Zealand context. We further categorised the relationship type as being adult, peer, university student, or mixed. Another important difference in our coding scheme was that we did not exclude mentoring studies based on the level of risk of the participants. Instead, we coded for the level of risk including: typical/low (typical or community youth with little or no risk and no non-normative problems), at-risk (youth with risk factors associated with poverty, school problems, family problems and/or low self-esteem, but no severe problems), high risk (youth who are offenders, have

substance problems, clinical/ mental health problems, severe family problems, or educational failure, and/or are living in an institutional setting) or mixed-risk (different youth in the programme have different levels of risk).

For the data extraction, the process involved a number of steps. The first five included studies were assessed by two independent coders (PI or SI and RA). After each one was coded, the two coders made comparisons. Any discrepancies in coding were discussed with the PI/SI. This consensus process was continued after every 5th study. The average inter-rater agreement was 80%. For the entire process, where there were discrepancies between coders, these were discussed and if agreement could not be made, the remaining researcher (SI or PI) entered into the discussion until consensus was reached.

A separate coding sheet was also created to assess the cultural appropriateness of the research and programmes. This was an important component as most of the research on youth mentoring is imported from the US; likewise, programmes are often developed using overseas models. In addition, as many programmes in New Zealand have been developed specifically for Māori and Pasifika youth, cultural coding was conducted from the Māori perspective, the Pasifika perspective, as well as from the overall New Zealand perspective. The cultural coding was completed by Māori and Pasifika representatives (Frank Solomon and Efeso Collins, respectively) of the research team when the programme had a significant proportion of relevant youth in their programme (i.e., 15% or more). Ann Dunphy coded for the overall cultural context, reflecting the ethnically-diverse youth population of New Zealand, if both Māori and Pasifika coding were not required. The coding sheet covered cultural appropriateness relating to the researcher(s), research procedures and analysis, programme implementers and programme design. Programmes were rated on the following scale:

- *Appropriate*: acknowledged cultural issues and demonstrated that they were taken into consideration; provided information about the researcher's and implementer's cultural backgrounds (cultural matching of either mentors or researchers was not required); included cultural competency training of the people delivering the programme; analysis took into account the cultural background of the participant; a recognition and understanding of culture;
- *Somewhat appropriate*: acknowledged cultural issues but did not significantly demonstration that these were taken into consideration;
- *Inappropriate*: acknowledged cultural issues but included an incorrect response to these issues; and
- *Ignored*: did not acknowledge cultural issues or programme implementers or researchers did not attempt to take them into consideration; no information provided about researchers or implementers.

For a number of studies, there was very limited information on programme characteristics. In cases where the study was evaluating a programme that fell under the umbrella of the Youth Mentoring Network, the Network contacted each individual programme to request additional information; this was then filled in on the code sheet. Coded data was then entered into Excel and imported into SPSS which was used to organise the analyses.

Assessment of methodological quality of included studies

Bias coding was conducted for each study. This involved assessing the bias of the research in regards to the methods, intervention, participation, and analysis (see bias coding sheet in Appendix A). Each study was assigned a score out of 25, which was then converted into a percentage, with articles with higher scores showing higher levels of bias. Articles were coded by two independent coders and the inter-rater agreement was 69%. Discrepancies between coders were discussed and when agreement could not be reached, the PI or SI made the final decision.

Measures of treatment effect

Treatment effects were measured in a number of ways as both quantitative and qualitative data were used in this review. Further, many of the studies that utilised quantitative data did not conduct statistical tests to determine effect and/or significance. As such, a variety of approaches were taken to determine the treatment effects. Individual effect indicators were determined for each outcome reported.

For quantitative studies that provided statistical results, an effect size (Cohen's *d*) for each measure was calculated (see data extraction coding sheet in Appendix A). If raw data were provided without any statistical tests, means and standard deviations were calculated from which effect sizes were calculated. Effect sizes of below .20 were seen as signifying unsuccessful outcomes; those with effect sizes between .20 and .35 were seen as indicating moderately successful outcomes, and effect sizes above .35 indicated successful outcomes. Once effect sizes were calculated, results were recoded for each goal domain (i.e., educational, psychological, behavioural, interpersonal, vocational, and cultural) as not effective, mixed or moderately effective, or effective. For a description of goal domains see Table 2. Mixed effects reflected multiple indicators within the same goal domain, but with inconsistent results. When programme studies indicated adverse effects, this was taken into consideration when determining the effectiveness of the programme. Programmes could be coded as effective in one goal domain and not effective in a different goal domain. Coding was done by both the PI and RA for all appropriate studies.

An effect size could not be calculated for all studies reporting quantitative data, for reasons such as raw data was not provided and only proportions were provided. Quantitative studies, for which an effect size could not be calculated, were coded for success based on the output they provided. The effects were coded using the same scale as above (i.e., not effective, mixed or moderately effective, or effective). This rating took into account the occurrence of any adverse outcomes. Again, if studies had more than one programme goal (e.g., educational and psychological) effectiveness was coded for each goal domain; all studies were coded by both the PI and the RA.

For qualitative studies, outcomes were coded in the data extraction code sheet for success, using the responses choices not effective, mixed results, and effective. To be coded as effective, all or most of the qualitative results needed to have indicated a positive effect. To be coded as mixed, some of the results needed to effective. To be coded as not effective, none or very few of the results were effective. Again, this rating took into account the occurrence of any adverse outcomes. As this was part of the data extraction, coding was completed by the RA and either the PI or SI. Once individual outcomes were assessed for effectiveness, this data was aggregated by domain taking into account adverse effects. Again, different goal domains could allow for different effectiveness ratings.

At the end of these processes, both quantitative and qualitative results were on the same scale which served two purposes: merging and comparing data. First, the merging of the types of data allowed for a more complete examination of the effectiveness of youth mentoring, as traditional systematic reviews typically focus on quantitative data only. Based on the combination of quantitative and qualitative results across goal domains, programmes were then coded for overall effectiveness in the following categories: not effective (not effective in any domain or using either methodology; very few effects found), mixed/moderately effective (effects found in some domains or had moderate effects across domains), effective (effective in many domains, possibly some minor variation by research methodology), or very effective (consistent, strong effects across domains and methodology).

The second purpose was to be able to compare the different methodologies. Comparing the methodologies allowed for both an examination in the variation of effectiveness as a result of method, as well as an examination of the quality of the research as a function of research method.

Results

The results are divided into three sections. The first section is a description of mentoring programmes in New Zealand that were included in this review. The second section is an analysis of the effectiveness of mentoring programmes, examining variation by programme characteristics. The third and final section is a description of the quality of the research on mentoring programmes. Table 3 provides a list of the programmes that are included in this review, indicating which are current programmes and members of the Youth Mentoring Network. A summary of the mentee/mentor, programme, and research characteristics for each study included in this review can be found in Tables 4 to 6.

Youth mentoring programmes in New Zealand

This review includes 26 evaluations covering 22 different mentoring programmes in New Zealand. (Note: one of these evaluations covered multiple programmes, but was assessed as a whole, as enough details of individual programmes were not provided). Of the programmes included in this review, 8 fall within the umbrella of the Youth Mentoring Network; and the majority of the remaining programmes (n = 9) were independent, one-off programmes, generally run in schools. According to the Network, there were 23 active programmes in New Zealand at the time of this review which meant that only 35% of the current programmes had evaluations of effectiveness for mentees completed by November 2009.

The following paragraphs describe overall characteristics of programmes included in this review. Due to the fact that not all evaluations provided full details on the programmes, percentages were calculated based on programmes with known characteristics and, therefore, were reported for valid cases only. For some programme features, there was a large amount of missing information. General programme characteristics tended to be more complete, whereas programme delivery information, such as the average length of the relationship, had substantial missing information.

The majority of programmes included in the review were independent (48%), with 20% being national programmes and 32% being regional multi-site programmes. The vast majority of programmes were located in urban areas (80%), with 8% being in rural areas and 12% being in mixed (both rural and urban) areas. Over half of the programmes were based at schools (65%) with some being based in the community such as at a centre or private facility (26%), and 9% mixed (both school and community/private facility based). Seventy percent of these programmes were established programmes, while 30% were relatively new, having been established in the 2 years previous to the research being conducted. The majority of programmes included in this review did not have a history of evaluation (70%); and all but one of the programmes were based on at least some principles of good practice.

From the studies included in this review, mentoring tended to be a component of a larger programme (64%), rather than a stand-alone programme by itself (36%). When mentoring was a component, the mentoring component made up varying proportions of the programme, with some programmes having the majority of services as mentoring (39%), while for others mentoring made up about a quarter (23%) or less than half (38%) of the programme. Of component programmes, apart from mentoring, the programme most often had an educational component (94%), or a life skills component (67%). Other components included interpersonal (50%), employment (31%), counselling (27%), behavioural (19%), and cultural (13%).

Mentoring tended to be one-on-one for most programmes (73%), with 12% of programmes having group mentoring and 15% of programmes having mixed mentoring. Eighty-three percent of programmes used established criteria for matching mentees and mentors, 56% of these programmes also matched on gender, and 21% of them also matched on ethnicity. For programmes, the most

common person who was the mentor was an adult (72%), with 12% of programmes having peer mentors, 12% using university students as mentors, and 4% using different types of mentors. In the majority of programmes (84%) mentors were volunteers, and most programmes screened (95%) and trained (91%) their mentors. In terms of support, once the mentorship had begun the largest proportion of programmes (61%) offered monthly ongoing support, while 22% of programmes offered this weekly, and 17% offered it every 2-4 months or less frequently. Further to this, 22% of programmes provided weekly supervision of the match, while 44% provided this monthly, and 33% provided it every 2-4 months or less frequently.

Most programmes (45%) tended to have some structure with 39% of programmes being highly structured and only 17% having little or no structure. Seventy-seven percent of programmes expected the mentors to meet weekly with mentees; while 18% expected contact every other week and only one programme expected contact less then every other month. There was a wide range in expected lengths of the relationship, from 2 to 48 months, across programmes, with an average expected length of 11.37 months (*SD* = 9.91). The most frequent expected length was 12 months, expected by 42% of programmes; 11% of programmes had expectations for the relationship to last longer than 12 months. Programmes tended to have minimal or no contact with families (71%), with the remainder having moderate to frequent family contact.

The majority of programmes targeted low Socio-Economic Status (SES) youth (44%), with 17% of programmes targeting mid SES youth and 39% of programmes targeting mixed SES youth. Most programmes also targeted at-risk youth (56%) or high-risk youth (20%). A further 18% provided mentoring to typical/low risk youth, with 8% of mixed risk status.

All but one programme (96%) identified at least one programme goal. Most programmes had more than one goal (i.e., 8% of programmes only had one goal, 40% had 2 goals, 12% had 3 goals, 20% had 4 goals, and 20% had 5 goals). Almost all programmes (96%) had educational goals. About half the programmes had psychological goals (52%), interpersonal goals (52%), behavioural goals (44%), vocational goals (40%), and a few programmes had cultural goals (20%).

Effectiveness of youth mentoring programmes

The effectiveness of youth mentoring programmes was first examined overall, regardless of goal domain (e.g., academic, psychological) or research methodology (i.e., quantitative or qualitative). Then effectiveness by domain and research methodology was examined. Next, based on the information available on each programme, features of programmes were examined to see if associations with effectiveness could be identified; these features included general programme characteristics, programme delivery characteristics, youth characteristics, and mentor-mentee relationships characteristics. Again, these findings are first presented as aggregated across goal domains and then within goal domains for educational, psychological, behavioural, and interpersonal. Vocational and cultural domains were not included as these were frequently not evaluated, even though they were programme goals 50% of the time for both goal-types (i.e., vocational and cultural). For the within-goal analyses, only findings that were different from the overall pattern are presented in order to minimise confusions by repeating findings. Research methodology was not examined at this level as it was a constant. Only characteristics that had adequate variability or had enough valid information are included. For instance, as 96% of programmes had identified goals, this variable was not examined as there was not enough variability to draw meaningful conclusions. Similarly, duration of the relationship was not included as only 23% included this in the research. **It is important to note that for the following results, patterns in the data are presented. This does not mean that if a particular type of programme was associated with less favourable outcomes that all programmes of that type were not effective, just that this was a trend. Also, a reminder is provided that statements made about the effectiveness of mentoring are tentative given the variability in the quality of the research included in this review (as discussed in detail below).**

Overall programme effectiveness

Overall, the majority of programmes included in this review showed some level of effectiveness (88%), with only 12% of programmes being classified as not effective. The remaining programmes were quite varied in their level of effectiveness: 35% were moderately effective (e.g., a modest effect size) or showed mixed effectiveness (e.g., effective in some goal domains but not others; effective for qualitative but not quantitative methodology); 27% were effective (e.g., strong effects found for many of the goals); and 27% were very effective (strong effects found for most or all of the goals). It should, however, be noted that for 26% of all programmes included in this review, at least one adverse outcome was found.

When examined by goal domain, programmes tended to be more effective in psychological and interpersonal areas (86% and 73%, respectively, of programmes were effective or very effective within these goals) and less so in academic, behavioural vocational, and cultural areas (45%, 36%, 40%, and 33%, respectively, of programmes were effective or very effective). This is, in part, because the effectiveness of programmes with academic and behavioural goals was very variable within programmes, meaning they were successful in some aspects of the goal but not others (e.g., attendance but not academic performance). In addition, programmes with academic goals were less consistent with their effectiveness than other types of goals, meaning there were programmes that were ineffective and programmes that were very effective. Another pattern that emerged in regards to programme goals was that programmes covering fewer areas, i.e., had fewer goal domains, tended to be more effective.

When examining the effectiveness of programmes, research methodology did seem to influence findings of effectiveness. Research designs utilising a qualitative methodology were more likely to indicate effectiveness than designs utilising a quantitative methodology. This pattern was found regardless of goal domain. This will be discussed in more detail in the following section on the quality of the research.

General programme characteristics

Concerning the mentoring programmes, general programme characteristics that were examined included: the dissemination of the programme (independent, multi-site regional, or national); programme age (new/relatively new versus more established); had a history of evaluation (yes or no); programme was based on principles of good practice (no, somewhat or yes); and programme was mentoring-only versus mentoring is one component of a wraparound programme. Overall, when looking at the dissemination of the programme, independent programmes tended to be less effective than regional multi-site programmes; there was not a trend for national programmes, meaning some were effective while others were not. One exception to this result was for behavioural goals where independent programmes showed a higher proportion of effectiveness (43%) as compared to regional multi-site or national programmes (both with 0% effective; 100% mixed/moderately effective).

For the age of the programme, those programmes which were more established were more likely to be effective as compared to new or relatively new programmes. This association was particularly true for programmes with interpersonal goals. Two thirds of new programmes were ineffective, whereas 80% of established programmes were effective.

Regarding a history of programme evaluation, more effective programmes tended to have a history of evaluation. Only unsuccessful programmes had not previously been evaluated. That said, many successful programmes had also not been previously evaluated.

There was also a positive association between best-practice and effectiveness. When programmes were based on international principles of best practice, they were more likely to be effective. This was

particularly true for psychological goals and for interpersonal goals. For psychological goals, for programmes that utilised best practices, all were successful, whereas programmes that somewhat utilised best-practice, 50% were ineffective or had mixed/moderate effects.

Interestingly, when mentoring was a component of a programme versus being a stand-alone programme, the programme was more effective. This said, it was not possible to tease apart the impact that mentoring had independent of the other programme components; therefore, this result should be interpreted with caution.

The only factor that did not have an association was the location of the programme. Variation in effectiveness was not found for school-based programmes or community-based programmes. There were too few mixed programmes to draw meaningful conclusions.

Programme delivery

The next set of factors examined was related to programme delivery. These included: type of relationship (one-to-one, group, or mixed); type of mentor (peer, university student, or other adult); compensation for mentors (paid versus volunteers); level of programme structure (in terms of time, location and activities); inclusion of families (no contact or at least minimal contact); expected duration of relationship (less than 6 months, 6 to 9 months, one year, or more than one year); and matching of mentors and mentees on gender or ethnicity.

Regarding the type of mentoring relationship, programmes where there was one-to-one mentoring or mixed (one-to-one and group) mentoring tended to be more effective than programmes that delivered mentoring in groups only. This was particularly evident for programmes with academic goals. Programmes that utilised peers as mentors were less effective than those with university students, but there was no clear pattern as to how the use of adults as mentors impacted on programme effectiveness.

There was also an association between level of structure and programme effectiveness. Programmes that were more highly structured, in terms of time, location and activities of mentoring, were more effective, as compared to programmes that had less structure. This pattern was found across programme goals.

However, programmes that had a longer expectation for the length of the mentee-mentor relationship tended to be more effective, once a minimal threshold was of more than one year was met. There was no consistency in effectiveness for programmes that expected the relationship to be a year or less.
There was no difference in effectiveness between programmes that paid mentors and those where mentors were volunteers. Likewise, programme effectiveness was not moderated by the inclusion of families in the programme.

Youth characteristics

In terms of youth characteristics, we examined variation in gender (male only, female only, or mixed); age as indicated by level of school they attend on average (primary, primary or intermediate, intermediate, intermediate or secondary, or secondary); socioeconomic background of the youth (low, mid, high or mixed); and risk status of the youth (low/no risk, at-risk, high-risk, or mixed). For gender, there was no variation in effectiveness between the programmes that were male only, female only or mixed males and females. Likewise for age and risk status, there was no variation in effectiveness by age or risk status.

When examining the socioeconomic background of the youth, there was a trend for programmes that were aimed at mid-level economic background youth being less effective as compared to programmes with low or mixed socioeconomic background youth. This variable, in particular, had high levels of missing information, so caution must be taken when interpreting this result.

Researcher-practitioner relations
In terms of the researcher-practitioner relationship, there was a moderate association between external assessment of programmes and greater programme effectiveness. This pattern was found consistently across programme goals, and particularly for educational goals. Related to this, in cases where the research evaluator also delivered the programme, there were slightly lower effectiveness rates seen.

Quality of research on youth mentoring
Of the programmes that were under the umbrella of the Youth Mentoring Network at the time of this review (n = 23), 35% have had some sort of research/evaluation on the effectiveness of the programme for the mentees. This means that most of the active programmes in New Zealand have not engaged in an evaluation of their programmes to assess effectiveness for their youth.

Of the studies included in this review, only 31% utilised a control group, and likewise only 31% utilised a pre-test post-test design. Very few studies (13%) assessed the effects of mentoring long term (i.e., 6 to 12 months post-intervention). Most (52%) assessed the effects either immediately post or within 2 months of the intervention, and 35% included only a mid-intervention assessment.

Methodology
In terms of research methodology and design, 31% of evaluations were purely quantitative studies, 31% were purely qualitative studies, and 38% were mixed studies, having both qualitative and quantitative components. No studies utilised an action research design. Interestingly, for mixed studies there were some discrepancies found between quantitative and qualitative outcomes. Of mixed studies, allowing for differing goal domains to be assessed independently, qualitative results showed more effective outcomes 60% of the time, quantitative results showed more effective outcomes for 7% of the time, and quantitative and qualitative methods reflected similar effectiveness 33% of the time.
One important indicator of research quality is passing the peer review process. Of the 26 studies included in this review, only 12% were peer reviewed.

Quantitative design
Looking at the quantitative studies, there was a range of different study designs seen in the literature included in this review. Experimental designs, of which there was only one in this review, were seen as the most rigorous of research designs. They were characterised by three key processes, the first being the presence of a control group, the second being the randomised allocation of participants to either the control or the treatment group, and the third being that there were multiple test-times during the research, i.e., a pre-test and post-test. This design is the only one that allows investigation into cause and effect (O'Leary, 2005).

Quasi-experimental design is a form of experimental research where there is no randomisation of groups. Quasi-experimental design are typically conducted with a control group, as was the case with 39% of the quantitative studies in this review, or sometimes without a control group, as was the case with 22% of the quantitative studies in this review. In quasi-experimental studies without a control, the treatment group is pre- and post-tested, allowing comparison between these results. Although these quasi-experimental designs permit some comparison and insight into the effectiveness of the intervention, they can not be seen to address cause and effect issues.

Lastly, correlational design, which 33% of the quantitative studies in this review used, does not include randomisation, does not have a control group and only involves post-testing of participants. Thus, though some relationship may be shown between outcomes and the intervention, statements about causation can not be made in the case of correlation studies. To draw dependable and accurate conclusions it is, therefore, necessary to use more experimental research designs.

Qualitative design

Of the studies included in this review, 18 included a qualitative research component, and 17% (n = 3) collected data both pre- and post intervention. Of these 18 studies, one was based on a single case study, 33% used multiple case studies, 78% face-to-face interviews, and 22% included focus groups. Fifty percent of these studies clearly stated the research paradigm used in the collection and interpretation of data. Most studies (67%) included statements indicating an understanding of the social context of the research participants. However, only five studies (23%) acknowledged reflexivity, an awareness of the effect the researcher may have on the construction and interpretation of results. While most studies (67%) provided an adequate description of the methods and procedures used regarding data collection, less than half included information on the data analysis process (39%). Further, only 39% of the studies showed evidence that their data analysis techniques were thorough and systematic. In terms of the interpretation and presentation of results, for most of the studies included in this review, the interpretation of results followed logically from the analysis (83%), and the findings appeared to match derived data (78%). Only 50% of studies noted limitations.

Bias

Given the low number of experimental studies in this review, it was important that all included studies were coded for bias. Bias coding involved assessing whether a study met standards for thorough and systematic research and, therefore, allowed us to gauge the quality of the research, and, consequently, the reliability of the findings being presented. The bias coding sheet examined aspects of bias relating to methods, intervention, participation, and analysis (see bias coding sheet in Appendix A). Each bias variable was coded as met or unmet and a total percentage bias was then calculated for each study. Of the studies included in this review, only one was classified as low bias, 31% were classified as having moderate bias, and 65% were classified as having high bias. This is not a surprising finding given the low use of experimental research designs. However, it is interesting to note that there was great variation in the types of bias seen across studies (see Table 7).

There were a number of particularly high areas of bias seen across the majority of studies in this review. Firstly, 58% of all studies showed high avoidance of performance bias, meaning that the intervention group received services beyond mentoring. This makes understanding the effectiveness of mentoring difficult, as it is hard to know whether outcomes are the result of mentoring, or due to the other services that were received. Secondly, half the studies showed strong attribution bias, where more than 20% of the participants dropped out of the programme. This could be problematic as those who left the programme may differ in characteristics to those remaining in the programme who are then assessed in the research. Regarding quantitative studies, there were some particularly high areas of bias seen. Eight-three percent of studies did not meet the selection bias, meaning they did not randomly assign participants to treatment or control groups, or they did not match the treatment and control group in the analysis. This is important as random assignment or matched-control accounts for a large amount of variance between groups based on group attributes; in the case where this is not met, differences between groups could be due to differing group attributes, rather than to the effect of the intervention. There was also high detection bias seen in quantitative studies in this review; in 79% of studies, the assessor was aware of the assigned treatment group of participants when collecting data. This could be a problem as there is a risk in the assessor unconsciously eliciting certain responses from the participants, particularly if the assessor has an investment in the programme, as was the case for a number of studies in this review.

For both quantitative and qualitative studies there was high bias regarding intention to treat, with 73% of quantitative studies and 57% of qualitative studies not including everyone who started the programme as potential participants in the research. Implications of those who had left or dropped out of the programme not being involved in the research means that results may not be accurate, as views and outcomes for those still in the programme and those who had left may be different. Half of all qualitative studies in this review also had high bias in terms of lacking thorough and systematic data

analysis techniques, and not having a clear theoretical framework to guide their data analysis. In qualitative research it is important to ensure that data are approached with a clear analysis framework and analysed in a thorough and systematic manner. Without this it is difficult to determine the extent to which the research results accurately reflect the collected data, ultimately impacting on the reliability of the results.

Cultural appropriateness

As previously discussed, given the unique cultural context of New Zealand, it was important that all studies were coded for cultural appropriateness. Cultural appropriateness coding involved assessing whether a study took into account the cultural needs of the targeted youth, particularly as this applied to Māori and Pasifika youth, as well as the overall New Zealand context (which includes Māori and Pasifika youth). As a reminder, the cultural appropriateness code sheet examined the cultural appropriateness of the researcher(s), research procedures and analysis, programme implementers, and programme design (see cultural appropriateness coding sheet in Appendix A). Coding was completed as appropriate for relevant research with relevant youth mentees; therefore, overall New Zealand culture was assessed for 18 studies, Māori culture was assessed for 14 studies, and Pasifika culture was assessed for 6 studies. Sum scores were calculated to indicate cultural appropriateness in the three cultural domains for research (research, procedures and analyses) and programme (implementer and design), separately.

For overall New Zealand culture, both programme and research were assessed to be highly appropriate. For programmes, only one was assessed as being moderately appropriate whereas all of the others (n = 17) were assessed as being highly appropriate. Likewise, for research, the same results were found with all research being assessed as highly appropriate with one being assessed as moderately appropriate. Overall, both programmes and those assessing the programmes were very responsive to the general New Zealand context.

For Māori cultural, the findings were less favourable. Examining programmes first, a large proportion (50%) was rated as having completely ignored Māori culture. Less than one quarter (21%) were rated as being highly culturally appropriate. Results were similar for research. Again, 50% of the research completely ignored Māori culture with only 29% being highly appropriate.

For Pasifika culture, findings, again, were less favourable, but slightly better than for Māori. For programmes, only 1 of the 6 was highly appropriate with most falling in the moderate range (n = 3, 50%). The remaining were low (n = 1) or ignored (n = 1). Slightly more favourable results were found for research, with 50% being highly appropriate (n = 3), 1 being moderately appropriate, and 2 had low levels of appropriateness.

Finally, programme cultural appropriateness was examined to see if there was an association with programme effectiveness. As there was little variability for overall New Zealand culture, this was not included in these analyses. For Māori cultural appropriateness, there appeared to be a negative association: programmes that were more culturally appropriate tended to be less effective. For instance, 3 of the 7 programmes that ignored Māori culture were effective (effective or very effective); whereas, none of the highly culturally appropriate programmes were effective (all were ineffective or mixed/moderately effective). For Pasifika cultural appropriateness, there did not appear to be an association with programme effectiveness. However, as the sample was relatively small, associations may not have been detected.

Discussion

This systematic review had four aims: to examine the effectiveness of youth mentoring programmes in New Zealand; to identify the characteristics of successful programmes; to assess the quality of the research on youth mentoring; and to identify gaps in the literature and recommend directions for future

research. Of the 13,292 unduplicated documents that were initially examined, 26 studies were ultimately retained, with a fairly even distribution of qualitative, quantitative and mixed-methods studies. This discussion will start by presenting the state of youth mentoring in New Zealand in general, followed by a discussion on characteristics of effective programmes and the quality of research on youth mentoring. It will then conclude with recommendations for programmes and researchers in the field of youth mentoring. It is important to note, that due to the great variability in the quality of the research, the findings presented here related to effectiveness should be interpreted as tentative. It is important to provide a reminder that most of the programmes evaluated in this review were not active programmes.

The state of youth mentoring in New Zealand

Youth mentoring has grown significantly in the past 20 years. There are now 23 active programmes within the Youth Mentoring Network umbrella. Of the programmes included in this review, most were independent, possibly reflecting the early stage that mentoring in New Zealand is in. Programmes have not had as much time as compared to other countries with large national programmes (e.g., Big Brothers/Big Sisters in the U.S. and Perach in Israel) to widely-disseminate their programmes. With time, it is anticipated that effective programmes will continue to grow in New Zealand and become more prevalent.

Interestingly, schools were the most popular location for the programmes. This is likely due, in part, to the finding that programmes with educational goals were the most prevalent with 94% having at least one. This reflects the very functional nature of mentoring programmes and the strong connection between the education and mentoring sectors. It also reflects the understanding of mentoring programmes that many at-risk youth leave school with little or no qualifications (MYD, 2003).

On a related note, few programmes incorporated cultural goals (20%), despite the fact that 62% (n = 16) had a significant proportion of Māori and/or Pasifika youth. This gap is particularly poignant given that cultural identity is an important component of well-being for Māori and Pasifika youth (Anae, 2001; Borell, 2005).

Most of the programmes included in this review utilised a traditional one-to-one mentoring relationship. No programmes utilised a team mentoring approach or e-mentoring. This prevalent use of the traditional model is despite the argument (see Evans and Ave, 2000 for an example) that this model may not be appropriate in the New Zealand context. Further, most programmes expected the mentors and mentees to have contact on a weekly basis. However, due to a lack of reporting, it is unclear what proportion of mentors and mentees actually met that expectation.

On a very positive note, almost all programmes screened their mentors and provided them with initial, and to a lesser degree, ongoing training or support and supervision. This is important as previous research has identified ongoing support as a characteristic of effective programmes (Bullen et al., in press). Clearly, presence of training is not equal to the quality of training. Therefore, it is unclear the degree of variation in the quality of training and mentor support.

There was variation in the cultural appropriateness of mentoring programmes. On one hand, programmes were highly appropriate for the overall New Zealand cultural context. This is important as much knowledge of mentoring is imported from overseas, particularly the U.S. As such, programmes need to ensure that they meet the needs of New Zealand youth by utilising an overall New Zealand cultural framework. On the other hand, programmes working with Māori and Pasifika youth were less culturally appropriate overall. This was particularly evident for programmes with Māori youth. In light of the significant proportion of Māori and Pasifika youth participating in youth mentoring programmes in New Zealand, it seems essential that established effective programmes take the opportunity to incorporate elements of Māori and Pasifika culture within their programme delivery. Programmes which are highly appropriate in terms of culture need to incorporate elements of best practice,

particularly around administration and quality assurance. It must be emphasised that the findings relate to mentoring only and do not reflect the effectiveness of other programmes working with Māori and Pasifika youth and their families.

Characteristics of effective programmes

As previously indicated, the varying quality of the research on mentoring programmes makes it difficult to draw strong conclusions about the characteristics that moderate programme effectiveness. That said, when placed within the context of international research, some conclusions can be drawn. To start with, programmes with psychological and interpersonal goals were typically more effective than programmes with other types of goals. There are two non-competing explanations for this finding which will focus on education, as the vast majority of programmes had at least one educational goal. First, it may be harder to make changes in the educational domain. This is likely to require more structure, intense focus and having the young person work on something that may not be as much fun, as youth with these goals have struggled in this area. Working on interpersonal and psychological goals may be more informal and more fun for the young person. This will make it easier to engage the young person and consequently may be more fun for the mentor. A second explanation is that a key premise of youth mentoring is that the mentor and mentee establish a high-quality relationship (Evans & Ave, 2000). The very nature of developing this relationship can serve to work on interpersonal and psychological goals, whereas, it is not automatic that it will support educational goals. It also appeared that programmes that direct their attention to fewer goals are more effective. It may be the case that programmes with educational goals either need to put high levels of structure and focus in place to specifically address these goals or provide other wraparound services such as tutoring to address these goals. In light of the evidence that many youth are underachieving at school (MYD, 2003), interventions such as youth mentoring need to help improve educational achievement. This discussion is not intended to minimise the importance of addressing psychological and interpersonal issues. As many New Zealand youth have experienced stress within the home (MYD, 2003), having a mentor may help them to cope better.

While component studies (i.e. those that were part of wraparound services) were shown to be more effective, it is difficult for this review to tease out the direct effects of mentoring as mentoring represented a component of most (64%) studies included in this review. Interestingly, previous international research (i.e., DuBois et al., 2002) did not find component studies to be more effective as compared to mentoring alone. One possible explanation for this is the difference in goals, in that New Zealand programmes, as previously mentioned, are highly focused on educational goals, whereas American programmes, the focus of the DuBois et al. (2002 meta-analysis, are less focused on educational goals. Regardless, it is likely that mentoring can provided an additional component to programmes that work with vulnerable youth.

It is interesting to note that so few studies (29%) included some level of family involvement as this appears to be salient within the cultural context of New Zealand, particularly for Māori and Pasifika youth (Anae, 2001; Hawk et al., 2001). While there did not appear to an association with programme effectiveness, there was relatively little variation on this construct as most programmes did not include the involvement of families. As international research (e.g., DuBois et al., 2002) has indicated that parent involvement is associated with more effective programmes, it seems worthwhile to examine this issue in greater detail.

One-off programmes, for the purpose of research (e.g., Master's thesis) or for single-school use, are largely ineffective. This is also probably reflected in there being a moderate association between external evaluation and greater programme effectiveness, as one-off programmes tend to be evaluated by the person administering the programme. Great caution should be taken by individuals who are considering this avenue of study or intervention as these programmes, generally-speaking, may not have the appropriate knowledge of best practice and the support and resources required to be effective. For those earning a Master's degree, it would be better to support an established

programme by providing an external evaluation. Those within schools, who are interested in starting a mentoring programme, should collaborate with existing programmes that have been shown to be effective.

In addition, programmes aimed at low or mixed economic background youth were more effective than programmes aimed at youth from mid-level economic backgrounds. Therefore, it is important for programmes to focus their resources on youth with greater levels of need.

Inconsistent with some international findings (DuBois et al., 2002), this review found a positive association between the expected duration of the mentor-mentee relationship, once a minimum threshold of one year was reached and the effectiveness of the programme. As there was large variation in expectations of the length of mentor-mentee relationships (i.e., 2-48 months), programmes should consider extending their programme as a means of increasing effectiveness.

Also inconsistent with the DuBois et al. (2002) meta-analysis is the lack of variation in effectiveness as function of risk status. DuBois and colleagues found that programmes that targeted at-risk youth were more effective than programmes that targeted youth who were not at-risk. One possible explanation for this difference relates to programme goals. In the New Zealand context, it is possible that programmes targeting typical youth may have goals that are appropriate for those youth (i.e., are not trying to effect substantial change where substantial change is not required); likewise, the outcomes may align well to the goals.

The final point under characteristics of effective programmes to be discussed is the importance of ensuring that principles of best practice are incorporated within mentoring programmes. While a number of programmes evaluated demonstrated these principles, a significant proportion did not, which was associated with less effective programmes. Best practice principles can provide a blue-print for programmes to follow and work towards meeting.

One particularly alarming finding was that there was a negative association between programme effectiveness and cultural appropriateness for Māori youth. One possible explanation is that programmes with Māori youth have typically followed one of two pathways that are at odds with each other. One pathway may have been to ensure that the programme is high in cultural appropriateness, but at the expense of fulfilling programme goals and incorporating principles of best practice. The other pathway may have been for programmes to ensure that programme goals are met and utilising principles of best practice at the expense of cultural appropriateness. The apparent clash between these two elements of programmes (effectiveness and cultural appropriateness) is likely to reflect the relative youth of the mentoring sector in New Zealand. They, therefore, do not need to be in opposition, and, in fact, should complement one another in strengthening mentoring programmes.

The quality of research on youth mentoring

Regarding the research on the mentoring programmes, there is concern that only 35% of current programmes had completed evaluations by November, 2009. Those programmes that have undergone and continue to undergo the evaluation process should be commended as it is a principle of good practice and was associated with effectiveness.

There is concern that the quality of the research, regardless of methodology, was very inconsistent with some high quality research, but also a fair amount of poor quality research. As a result of this variability, it makes conclusions about effectiveness much less reliable and harder to draw. One example of where this was seen was the large proportion of studies with high avoidance performance bias (57%), indicating that most evaluations did not include participants who did not complete the programme, potentially inflating effectiveness.

Another important area of consideration of the research is that very few studies (13%) included assessment of effectiveness in the long-term (i.e., more that 6 months post-intervention), with most (52%) measuring effectiveness in the short-term (i.e., up to two months post-intervention). This makes it difficult to know if there were any lasting effects of the programme.

One interesting finding was that the effectiveness of programmes varied as a function of methodology. This difference has two possible explanations. First, it could be said that qualitative approaches are more sensitive to effects, particularly for studies with small sample sizes. However, it could also mean that the qualitative methods used were more biased and reported effects may not have been accurate. Future research should clarify this issue.

An additional point is the overall lack of consideration of culturally appropriate research methods when evaluating programmes with Māori and Pasifika youth. It should be noted that some research was culturally appropriate, but this was a relatively small proportion. By not utilising a cultural framework, it is possible that effects were not found that may be present.

The final point of discussion regarding research on youth mentoring is that this review focused on the effectiveness of youth mentoring, not the cost-effectiveness of youth mentoring. Cost effectiveness analysis reflects how much a unit of achieved outcome costs so that comparisons can be made about the efficiency of programmes. While this is a very important assessment, it is beyond the scope of this review.

Conclusions and Recommendations

Youth mentoring has clearly grown in scope over the past 20 years. There are now 23 active programmes under the umbrella of the Youth Mentoring Network, from all over the country, reflecting the broad uptake of this effective (DuBois et al., 2002), social intervention for young people. While still relatively young, the youth mentoring movement has the opportunity to create very effective programmes by incorporating principles of best practice identified in this review and placed in the context of international research. It is essential that mentoring continue to improve upon its practice as the concept of mentoring fits into the PYD framework (e.g., Farruggia & Bullen, in press; Larson, 2000) which emphasises that every young person has potential.

Recommendations for programmes:

- Programmes need to be evaluated for effectiveness. Only a small proportion (35%) of known, active programmes (n = 23) have had any evaluations on the effectiveness of their programmes for mentees, regardless of the quality. Funding needs to be made available to programmes so that they can engage in this type of work;
- Programmes that are ineffective or have mixed results should ensure that they incorporate principles of best practice within programme delivery;
- Programmes that are ineffective or had mixed results with large numbers of programme goals should consider becoming more specialised, focusing on fewer programme goals;
- Programmes need to consider how culture provides an important context for many youth:
    - Effective, but less culturally appropriate, programmes working with Māori and Pasifika youth may increase their effectiveness and reach by incorporating a cultural framework within their programme delivery;
    - Less effective, but culturally appropriate, programmes working Māori and Pasifika should incorporate elements of best practice, particularly around administration and quality-assurance; and

- Programmes working with Māori and Pasifika youth should consider adding cultural goals (if appropriate for their programme).

Recommendations for researchers:

- Researchers need to eliminate the high levels of bias found in much of the research currently conducted on youth mentoring in New Zealand. Highlighted recommendations include:
  - using control groups, preferably with random assignment or at least matched comparison;
  - including youth in the evaluation who leave the programme;
  - conducting pre-tests;
  - conducting post-tests with adequate time between programme close and the assessment (e.g., 6 months or more);
  - submitting the results of their evaluations for peer reviewed publications; and
  - engaging external programme evaluators.
- Research needs to address all goals of the programme, only 50% of evaluations covered all programme goals.
- Research reports need to provide adequate detail on:
  - their framework, methods, analyses and results to accurately communicate the results of the study and to demonstrate the quality of the research conducted; and
  - the programme delivery.
- Research needs to utilise a cultural framework when evaluating programmes with significant proportions of Māori and Pasifika youth.
- Using good research procedures, future research should examine:
  - how variation in the quality of training may be associated with programme effectiveness;
  - the strength and quality of the relationship between mentors and mentees;
  - if parental/family involvement in mentoring is associated with more effective programmes, and if there is variation by type of goal;
  - how different research methodologies are associated with effectiveness and establish principles of best research so that quantitative and qualitative methodologies can both be used with certainty that the findings are accurate; and
  - if effectiveness is increased for programmes by incorporating cultural goals into the programme.
- Research needs to examine the cost-effectiveness of youth mentoring.
- This review should be updated in five years and only include studies that have less bias. In addition, current research questions should be re-examined to ensure that the findings are replicated.

Acknowledgements

The review was conducted with support from the Ministry of Youth Development and Health Research Council.

Potential conflicts of interest

Efeso Collins is a Programme Director of Project K and a personal friend of the Founder and General Manager of Affirming Works; he previously held a position on the MATES Advisory Board. Ann Dunphy is a current member of the MATES Advisory Board.

References

Included studies

Adams, R. J. (2004). *Constructing meaning from mentoring: The experiences of mentors and mentees*. Unpublished master's thesis, University of Canterbury, Christchurch, New Zealand.

Afeaki-Mafile'o, E. (2007). Affirming Works: A collective model of Pacifika mentoring. In P. Culbertson, M. Nelson Agee, & C. 'Ofa makasiale (Eds.), *Penina uliuli: Contemporary challenges in mental health for Pacific peoples* (pp. 16-25). Honolulu, HI: University of Hawai'i Press.

Ave, K., Evans, I., Hamerton, H., Melville, L., Moeke-Pickering, T., & Robertson, N. (1999). *Mentoring for children/youth at risk demonstration project: Final evaluation report*. Wellington, New Zealand: The Crime Prevention Unit, Department of the Prime Minister and Cabinet.

Ballinger, B., Mason, N., & Waring, M. (2009). YWCA Future Leaders Evaluation: A report for the YWCA Auckland. Auckland, New Zealand: AUT University, Consulting and Training.

Deane, K. (2009). *Individual growth in a community-based youth development program*. Auckland, New Zealand: The University of Auckland, Department of Psychology.

Deane, K. and N. Harre (2008). *Executive summary of results for Project K quantitative evaluation measures from pre- to post- programme.* Auckland, New Zealand: The University of Auckland, Department of Psychology.

Enkey, R. F. (2001). *A Research and Programme Evaluation of the North Shore Education Trust Project K*. Auckland, New Zealand: Project K.

Hammond, J. (2007). *Relationships matter: An evaluation of the Student Engagement Initiative at a girls' secondary schoo*l. Unpublished research project, University of Canterbury, Christchurch, New Zealand.

Heke, J. I. (2005). *Hokowhitu: A sport-based programme to improve academic, career, and drug and alcohol awareness in adolescent Māori*. Unpublished doctoral thesis, University of Otago, Dunedin, New Zealand.

Hill, J. (2008). *'I Have A Dream' A report for the IHAD Foundation: An evaluation of the programme and outcomes for students*. Auckland, New Zealand: The Education Group Ltd.

Irving, E., Moore, D. W., & Hamilton, J. (2003). Mentoring for high ability high school students. *Education & Training, 45*, 100-109.

Kostuk-Warren, J. (2005). *Adolescent well-being: Effects of time and intervention. Psychology*. Unpublished doctoral thesis, The University of Auckland, Auckland, New Zealand.

Lennan, M. (2006). *Evaluation report of the Big Buddy Mentoring Trust*, Auckland, New Zealand.

Litchfield, J. E. (2006). *Boys talk about mentoring and "stuff": A small-scale evaluation of Y9-Y13 mentoring relationships in a boys' high school*. Unpublished master's thesis, The University of Auckland, Auckland, New Zealand.

Lyon, D. R. (1992). *The adjustment problems of Cambodian secondary school students: An exploratory survey and an innovative buddy system intervention*. Unpublished master's thesis, The University of Auckland, Auckland, New Zealand.

McClean, D. (2007). *Evaluating the early impact of the Rangatahi Māori Mentoring Program developed to encourage and support Māori secondary school students to pursue a career in health*. Unpublished master's thesis, Griffith University, Queensland, Australia.

McInerny, J. (2005). *The Buddy Programme at Presbyterian Support Southland Child and Family Services: Five years on….* Unpublished master's thesis, University of Otago, Dunedin, New Zealand.

Milne, B. J., Chalmers, S., Waldie, K. E., Darling, H., & Poulton, R. (2002). Effectiveness of a Community-based Truancy Intervention: A pilot study. *New Zealand Journal of Educational Studies, 37*, 191-203.

Ministry of Education (2009). *Review of targeted policies and programmes: Mapihi Pounamu and He Ara Tika Maori participation initiatives*. Wellington, New Zealand: Author.

Qiao, C., & McNaught, H. (2007). *Evaluation of Project K*. Wellington, New Zealand: Ministry of Social Development , Centre for Social Research and Evaluation Te Pokapū Rangahau Arotaki Hapori.

Selwood, J. (2005). *Perspectives on a deaf mentoring programme: Does it make a difference?* Unpublished research project, Christchurch College of Education Christchurch, New Zealand.

Starpath (2006). *MATES 2005 Evaluation*. Starpath technical Report 15.1. Auckland, New Zealand: Author.

Starpath (2007). *MATES 2006 Evaluation*. Starpath Technical Report 15.2. Auckland, New Zealand: Author.

Stevenson, L. (2008). *Personalised learning in a web 2.0 environment*. Unpublished master's thesis, University of Waikato, Hamilton, New Zealand.

Tasi, B. S. (2009). *Supporting youth for work in New Zealand: A case study of the Samoan experience*. Unpublished master's thesis, The University of Auckland, Auckland, New Zealand.

Wilson, S. (2006). Improving retention and success: A case study approach for practical results. *Journal of College Student Retention: Research, Theory and Practice, 7*, 245-261.

Youth at Risk of Offending Team. (2001). *Evaluation of the one to one mentoring programme, Nelson*. Nelson, New Zealand: Author.

Excluded studies

Aronson, C. (2007, June). The young and the refuse at THE EDGE. *The Big Idea: Te Aria Nui*.

Baker, W., & McNicoll, A. (2006, July). *I can get by with a little help from my friends: Peer mentoring - critical friends for the reflective practitioner*. Paper presented at the 2006 Annual International Conference of the Higher Education Research and Development Society of Australasia Inc. (HERDSA), The University of Western Australia, Australia.

Blechman, E. (1991). *Mentors for high-risk minority children: Toward bicultural competence. Mentoring programs for young minority males*. Unpublished manuscript, University of Colorado, Boulder, CO.

Bolton, J. M. (2007). Developing composition pedagogical knowledge: Music teacher education students as online mentors. *Australian Online Journal of Arts Education, 3*, 1-13.

Brown, M. (2004). Notschool.net uncovered: Not all it seems. *New Zealand Journal of Teachers' Work, 1*, 7-15.

Campbell, T. (2007). *"New Traditions": Best practice for promoting Maori secondary school students into health careers*. Auckland, New Zealand: Auckland District Health Board.

Cardy, T. (1999, July 7). Mentors set a good example. *Dominion Post*, p.7.

Clarke, E. M. (1998). *Te wananga o te hinengaro: A critique of the pedagogical context of Maori mentoring*. Unpublished master's thesis, The University of Auckland, Auckland, New Zealand.

Collins, S. (2008, February, 9). Being there for the teens in trouble: Making a difference no simple matter. *New Zealand Herald*.

Cox, R. (1999). I've got a mentor: Teenagers on-track through help programme. *New Zealand Principal, 14*, 12-14.

Cox, R. (2002). *Report on New Zealand secondary schools pastoral care survey*. Auckland, New Zealand: Project K.

Cox, R. (2004). Sowing the seeds of mentoring. *New Zealand Education Review, 9*, 6-7.

Davies, A., Lowry, J., McBride P., Desai, S., Withers, V., & Sundgren, W. (2009). *Youth mentoring relationships: A youth development wrap-around approach. Literature Review of better practice*. Unpublished manuscript, Youthline, Auckland, New Zealand.

Deane, K. H., & Moore, J. (2009). *Individual growth in a community-based youth development program*. Auckland, New Zealand: The University of Auckland, Department of Psychology.

Denny, S. (2004). *Healthy youth development: A review of effective programs and approaches*. Auckland, New Zealand: Waitakere Education Sector Trust.

Densem, P., & Beard, L. (1984). *Nurturing academic talent in schools: The development and implementation of the Christchurch Learning Activity Mentor Programme.* Paper presented at the Third National Conference on Gifted and Talented Children, Christchurch, New Zealand.

Donnelly, F. (1961). *A youth club in a changing town a study of an effort made in a changing old mining town to assist the development of the youth of the area through a youth club*. Unpublished research project, University of New Zealand.

Du Chateau, C. (2006, August 12). The power of one keen mentor. *New Zealand Herald*, B5.

Eller, M. (2004). *Moving on up: Mentoring and the school to work transitions of young people in NZ*. Unpublished manuscript.

Elliott, K. J., & Lambourn, A. J. (1999). Sex, drugs and alcohol: Two peer-led approaches in Tamaki Makaurau/Auckland, Aotearoa/New Zealand. *Journal of Adolescence, 22*, 503-13.

Evans, I. M., & Ave, K. T. (2000). Mentoring children and youth: Principles, issues and policy implications for community programmes in New Zealand. *New Zealand Journal of Psychology, 29*, 41-49.

Farruggia, S., Bullen, P., & Van Ross, R. (2009). *Progress Report: MATES Junior Programme Evaluation*. Auckland, New Zealand: The University of Auckland.

Gearing, N. (1999). Brothers and sisters. *Listener, 168*, 32-33.

Gilmour, J. A., Kopeikin, A., & Douche, J. (2007). Student nurses as peer-mentors: collegiality in practice. *Nurse Education in Practice, 7*, 36-43.

Hill, J. (2005). *'I Have A Dream' The Third Formative Evaluation of the Programme*. Mt Roskill, Auckland, New Zealand: IHAD Foundation.

Hill, J. (2006). *'I Have A Dream' The Forth Formative Evaluation of the Programme*. Mt Roskill, Auckland, New Zealand: IHAD Foundation.

Hill, J. (2007). *'I Have A Dream' A Report for the IHAD Foundation: An Evaluation of the Programme and Outcomes for Students*. Mt Roskill, Auckland, New Zealand: IHAD Foundation.

Holland, J., Meikle, H., Ellison, K., & Allan, A. (2008). Inspiring the talented: A New Zealand high school and university team up. *TalentED, 25*, 54-58.

Hynes, K. (2006). *20 years on: Evaluation of the peer support programme in New Zealand secondary schools*. Christchurch, New Zealand: Peer Support Programme New Zealand.

Isles, P., & Freer, R. (1999). Students helping students. *Nursing New Zealand, 5*, 18-9.

Joseph, R. (2007). Whanau mentoring, Maori youth and crime: Possible ways forward. *Childrenz Issues: Journal of the Children's Issues Centre, 11*, 26-32.

Maere, C. (2009). *Best practice when mentoring young offenders*. Wellington, New Zealand: Ministry of Justice.

Middleton, S. (2007). Mentoring and teaching in academic settings: Professional and cultural identities from one Pākehā's perspective. *MAI Review, 3*.

Milardo, R. M. (2005). Generative uncle and nephew relationships. *Journal of Marriage and Family, 67*, 1226-1236.

Moore, J. (2005). *Self-efficacy and health behaviours: A test of measures to assess the effectiveness of a youth development programme*. Unpublished master's thesis, The University of Auckland, Auckland, New Zealand.

Morgan, B. (2006). *Nga hoa o te kupenga rorohiko: An innovative distance learning programme to address the paucity of Māori engineering enrolments*. Proceedings of the Australasian Association for Engineering Education creativity, challenge, change: partnerships in engineering education conference. Auckland, New Zealand.

O'Neill, M. (2005). *A mentor training programme: Evaluation of effectiveness*. Unpublished master's thesis, The University of Auckland, Auckland, New Zealand.

Park, H. (1999). *Te Timatanga Hou: The New Beginning*. Workshop paper presented at the Work Indigenous People's Conference, Albuquerque, NM.

Peters, J. (2000). *Models of Mentoring*. Wellington, New Zealand: Ministry of Education.

Salmond, A. (2003). *The Starpath Project for Participation and Success. Proposal for Partnership for Excellence*. Auckland, New Zealand: The University of Auckland, Starpath.

Shepheard, N. (2005). Unfurling the koru: 10 years of Project K. *North and South, 235*, 92-101.

Smith, A. (2004). Project K: Experiential learning at heart of success. *People and Performance, 12*, 6-10.

Tangi-Metua Williams, B. (2005). *The gift of dreams: Effective mentoring for Pacific secondary school students*. Unpublished master's thesis, The University of Auckland, Auckland, New Zealand.

TYLA. (2001). *Evaluation of TYLA, Avondale*. Auckland, New Zealand: Turn Your Life Around (TYLA).

Velde, M. (2000). *Dealing with drugs: Limits and precautions*. New Zealand Education Gazette, 79, 4-10.

Worrall, J. (2003). *TYLA Turn Your Life Around: A police school-based early intervention programme: A process evaluation*. Auckland, New Zealand: Turn Your Life Around (TYLA).

Young, T. (2007). E*valuation of Community Youth Programmes: Summary Report*. Wellington, New Zealand: Ministry of Justice.

Zakharov, K., Mitrovic, A., & Johnston, L. (2007). *Pedagogical Agents Trying on a Caring Mentor Role*. Christchurch, New Zealand: University of Canterbury.

Studies awaiting assessment

Beard, E. M., & Densem, P. A. (1986). News around the World. *Gifted Education International, 4*, 111-15.

Mitchell, S. A. (2006). *SPIES: An evaluation of tutoring programs for Maori and Pasifika students at the Faculty of Engineering, University of Auckland*. Proceedings of the Australasian Association for Engineering Education creativity, challenge, change: Partnerships in engineering education conference. Auckland, New Zealand.

Ongoing studies

Gordon, L. (2009). Invisible children: First year research report 'A study of the children of prisoners'. Christchurch, PILLARS Inc.

Farruggia, S., & Bullen, P. (2009). MATES Junior Programme Evaluation: Final report. Auckland, New Zealand: The University of Auckland.

Other references

Adolescent Health Research Group. (2008). *Youth'07: The health and wellbeing of secondary school students in New Zealand: Initial findings*. Auckland: University of Auckland.

Allen, T. G., Rhodes, J. E., & Eby, L. T. (2007). Definition and Evolution of Mentoring. In T. G. Allen & L. T. Eby (Eds.), *The Blackwell Handbook of Mentoring: A Multiple Perspectives Approach.* Malden, MA; Oxford: Blackwell Publishing.

Anae, M. (2001). The new 'Vikings of the sunrise': New Zealand-borns in the information age. In C. Macpherson, P. Spoonley & M. Anae (Eds.), *Tangata O Te Moana Nui: The Evolving Identities of Pacific Peoples in Aotearoa/New Zealand* (pp. 101-121). Palmerston North, New Zealand: Dunmore Press Ltd.

Auckland Youth Mentoring Association (2002). *Youth mentoring: An advice manual for Manukau and beyond.* Auckland, New Zealand: City of Manukau Educational Trust.

Borell, B. (2005). Living in the city ain't so bad: Cultural identity for young Maori in South Auckland. In J. H. Lui, T. McCreaner, T. McIntosh & T. Teaiwa (Eds.), *New Zealand identities: Departures and destinations* (pp. 191-206). Wellington: Victoria University Press.

Bullen, P., Farruggia, S. P., Rozas Gómez, C., Hebaishi Hasan Kamal, G., & Mahmood, M. (in press). Meeting the Graduating Teacher Standards: The added benefits for university students who mentor youth.

Clark, T. C., Robinson, E., Crengle, S., Grant, S., Galbreath, R. A., & Sykora, J. (2009). Youth'07: The Health and Wellbeing of Secondary School Students in New Zealand. *Findings on Young People and Violence*. Auckland: The University of Auckland.

DuBois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. *American Journal of Community Psychology, 30*, 157-197.

DuBois, D. L., & Karcher, M. J. (2005). Youth mentoring: Theory, research and practice. In D. L. DuBois & M. J. Karcher (Eds.), *Handbook of youth mentoring* (pp. 2-12). Thousand Oaks, CA: Sage Publications.

Durie, M.H. (1998). *Whaiora: Mäori health development.* Auckland: Oxford University Press.

Evans, I. M., & Ave, K. T. (2000). Mentoring children and youth: Principles, issues, and policy implications for community programmes in New Zealand. *New Zealand Journal of Psychology, 29*, 41-49.

Farruggia, S.P., & Bullen, P. (in press). Positive Youth Development in Aotearoa/New Zealand. In J. Low & P. Jose (Eds.). *Lifespan Development: New Zealand Perspectives* (2nd Edition).

Farruggia, S., Bullen P., & Davidson, J. (under review). Important non parental adults as an academic resource for youth.

Grossman, J. B., & Tierney, J. P. (1998). Does mentoring work? An impact study of the Big Brothers Big Sisters program. *Evaluation Review, 22*, 403-426.

Hammond, J. (2007). *Relationships matter: An evaluation of the Student Engagement Initiative at a girl's secondary schoo*l. Unpublished research project, University of Canterbury, Christchurch, New Zealand.

Larson, R. (2000). Toward a psychology of positive youth development. *American Psychologist, 55*, 170-183.

Lerner, R. M., & Castellino, D. R. (2002). Contemporary developmental theory and adolescence: Developmental systems and applied developmental science. *Journal of Adolescent Health, 31*, 122-135.

Littell, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic Reviews and Meta-Analysis*. New York, NY: Oxford University Press.

LoSciuto, L., Rajala, A. K., Townsend, T. N., & Taylor, A. S. (1996). An outcome evaluation of Across Ages: An intergenerational mentoring approach to drug prevention. *Journal of Adolescent Research, 11*, 116-129.

MENTOR/ National Mentoring Partnership (2005). *The elements of effective practice*. Retrieved from http://www.mentoring.org/program_staff/files/elementseffectivepractice.pdf.

Ministry of Education. (2006). *Student outcome overview 2001 – 2005*. Wellington, New Zealand: Author.

Ministry of Mäori Affairs/Te Puni Kokiri. (2000). *Progress towards closing social and economic gaps between Mäori and non-Mäori*. Wellington, New Zealand: Author.

Ministry of Pacific Island Affairs. (2003). *Ala fou - new pathways: Strategic directions for Pacific youth in New Zealand*. Wellington, New Zealand: Author.

Ministy of Youth Development. (2003). *12 to 24: Young people in New Zealand*. Wellington, New Zealand: Author.

Ministry of Youth Development. (2002). *Youth development strategy Aotearoa*. Wellington: Author.

O'Leary, Z. (2005). *Researching real-world problems: A guide to methods of inquiry*. Cornwall: Sage Publications.

O'Regan, T. (1987). Who owns the past? Change in Mäori perceptions of the past. In J. Wilson (Ed.), *From the beginning: The Archaeology of the Mäori* (pp.21-26). Dunedin: Otago University.

Pere, R. (1984). Te Oranaga o te Whanau. *Hui Whakaoranga: Mäori health planning workshop.* Wellington: Department of Health.

Search Institute. (2003). 40 developmental assets for adolescents (ages 12-18). Retrieved 26 June, 2008, from http://www.search-institute.org/system/files/40Assets.pdf.

Statistics New Zealand/Tatauranga Aotearoa. (2006). 2006 Census Data: Quick statistics about New Zealand. Retrieved 26 May 2008, 2008, from Statistics New Zealnd Website: http://www.stats.govt.nz/census/census.

Tolan P., Henry, D., Schoeny, M., & Bass, A. (2008). Mentoring interventions to affect juvenile delinquency and associated problems. Chicago: The University of Illinois at Chicago, Institute for Juvenile Research.

Werner, E.E., & Smith, R.S. (1982). *Vulnerable but not invincible: A longitudinal study of resilient children and youth*. New York: McGraw Hill.

Youth Mentoring Network (2009). *Guide to Effective Practice in Youth Mentoring New Zealand*. Auckland, New Zealand: Ministry of Youth Development.

Tables

Table 1 Excluded studies and reason for exclusion.

Citation	Reason for exclusion
Aronson, C. (2007).	Does not examine effectiveness of a programme
Baker, W., & McNicoll, A. (2006).	Does not examine effectiveness of a programme
Blechman, E. (1991).	Programme not set in New Zealand; does not examine effectiveness of a programme
Bolton, J. M. (2007).	Investigates effects on mentors not mentees
Brown, M. (2004).	Does not examine effectiveness of a programme; does not involve a formal mentoring programme; does not meet age criteria
Campbell, T. (2007).	Does not examine effectiveness of a programme
Cardy, T. (1999).	Does not examine effectiveness of a programme
Clarke, E. M. (1998).	Does not examine effectiveness of a programme; does not involve formal mentoring
Collins, S. (2008).	Does not examine effectiveness of a programme
Cox, R. (1999).	Does not examine effectiveness of a programme
Cox, R. (2002).	Does not examine effectiveness of a programme
Cox, R. (2004).	Does not examine effectiveness of a programme
Davies, A. et al., (2009).	Does not examine effectiveness of a programme
Deane, K. H., & Moore, J. (2009).	Data already used in included study
Denny, S. (2004).	Does not examine effectiveness of a programme; does not involve formal mentoring
Densem, P., & Beard, L. (1984).	Does not examine effectiveness of a programme
Donnelly, F. (1961).	Does not examine effectiveness of a programme; does not involve formal mentoring
Du Chateau, C. (2006).	Does not examine effectiveness of a programme
Eller, M. (2004).	Does not examine effectiveness of a programme; does not meet age criteria

Elliott, K. J., & Lambourn, A. J. (1999).	Does not examine the effectiveness of a programme
Evans, I. M., & Ave, K. T. (2000).	Does not examine effectiveness of a programme
Farruggia, S. P. et al., (2009).	Does not examine effectiveness of a programme; examines programme process
Gearing, N. (1999).	Does not examine effectiveness of a programme
Gilmour, J. A. et al., (2007).	Does not meet age criteria
Hill, J. (2005).	Same sample investigated as 2008 article included
Hill, J. (2006).	Same sample investigated as 2008 article included
Hill, J. (2007).	Same sample investigated as 2008 article included
Holland, J. et al., (2008).	Does not examine effectiveness of a programme
Hynes, K. (2006)	Does not examine effectiveness of a programme
Isles, P., & Freer, R. (1999).	Does not examine effectiveness of a programme; does not meet age criteria
Joseph, R. (2007).	Does not examine effectiveness of a programme; does not involve a formal mentoring programme
Maere, C. (2009).	Does not examine effectiveness of a programme
Middleton, S. (2007).	Does not examine effectiveness of programme; does not involve formal mentoring; does not meet age criteria
Milardo, R. M. (2005).	Does not examine effectiveness of a programme; does not examine formal mentoring
Moore, J. (2005).	Does not examine effectiveness of a programme
Morgan, B. (2006).	Does not examine effectiveness of a programme; does not meet the age criteria
O'Neill, M. (2005).	Examines effectiveness of a mentor training programme; does not examine effectiveness of programme for mentees
Park, H. (1999).	Does not examine effectiveness of a programme
Peters, J. (2000).	Does not examine effectiveness of a programme
Salmond, A. (2003).	Does not examine effectiveness of a programme
Shepheard, N. (2005).	Does not examine effectiveness of a programme
Smith, A. (2004).	Does not examine effectiveness of a programme
Tangi-Metua Williams, B. (2005).	Does not examine effectiveness of a programme
TYLA (2001).	Does not examine effectiveness of a programme; does not involve formal mentoring
Velde, M. (2000).	Does not examine effectiveness of a programme
Worrall, J. (2003).	Does not examine effectiveness of a programme; does not involve formal mentoring
Young, T. (2007)	Does not examine effectiveness of programme; does not involve formal mentoring
Zakharov et al., (2007)	Does not examine effectiveness of programme; does not meet age criteria

Table 2. Programme goals and specific outcome indicators.

Programme Goals	Outcome Indicators	
Educational	Academic problems	
		Truancy rates
		School (YSR*)
		Attendance/truancy
	Academic achievement	
		Enrolment in university
		Grades
		Graduation
		Homework
		Retention
	Academic adjustment	
		Academic aspirations
		Academic decision making
		Academic self-efficacy
		Attitude towards school
		Feeling about attending school
		Increased engagement
		Motivation towards school work
		On-task behaviour
		Requesting help from teachers
		Sense of school belonging
Psychological	Assertiveness	
	Confidence	
	Depression (YSR)	
	Help-seeking beliefs	
	Identity (YSR)	
	Mood swings	
	Personal development	
	Self-esteem	
	Self-efficacy	
	Somatisation (YSR)	
	Subjective well-being	
Behavioural	Social contexts	
		Aggression (YSR)
		Behaviour at home
		Behaviour at school
		Discipline problems
		Managing anger
		Prosocial activities
		Social presentation
	Health promoting behaviours	
		Activity (YSR)
		Physical activity
	Delinquency	
		Offending rates
		Delinquency (YSR)
		Detentions
	Substance use	

		Alcohol use Binge drinking Smoking Substance use
Interpersonal	Family relationships	
		Cohesion (FES^{+}) Conflict (FES) Expressiveness (FES) Family cohesion Home problems Parental monitoring
	Peer relationships	
		Communication with peers Relationship with friends Socialising outside school hours

	Other interpersonal	Approaching others socially Avoidance of communication Communication Leadership skills Social competence Social self-efficacy Social (YSR) Unpopularity (YSR)
Vocational	Career decision self-efficacy Discussions with parents about career	
Cultural	Identity Socio-cultural problems Use of NZ sign language	

*YSR = Youth Self Report (Achenbach & Edelbrock, 1987)
[+]FES = Family Environment Scales (Moos and Moos, 1981)

Table 3. Included studies, their programme name, current status, and if they are current and part of Youth Mentoring Network.

Authors	Programme name	Current programme	Part of Youth Mentoring Network
Adams, R. J. (2004)	District Truancy Services Mentoring Programme	No	No
Afeaki-Mafile'o, E. (2007)	Affirming Works	Yes	Yes
Ave, K. et al., (1999)	Multiple programmes: Man Alive, Just us Youth, Te Whare Putea, Presbyterian Support Services, Tuakana/Taina Mentoring Programme INC, Te Runanga O Ngati Porou	No	No
Ballinger, B. et al., (2009)	YWCA Future Leaders Evaluation	Yes	Yes
Deane, K., & Harre, N. (2008)	Project K	Yes	Yes
Enkey, R. F. (2001)	Project K	Yes	Yes
Hammond, J. (2007)	Student Engagement Initiative	No	No
Heke, J (2005)	Hokowhitu	No	No
Hill, J. (2008)	I Have a Dream Foundation	Yes	Yes
Irving, E. et al., (2003)	Mentoring programme for high ability high-school students	No	No
Kostuk-Warren, J. (2005)	Project K	Yes	Yes
Lennan, M. (2006)	The Big Buddy Mentoring Trust	Yes	Yes
Litchfield, J. E. (2006)	School based mentoring at a boys high school	No	No
Lyon, D. R. (1992)	Buddy System (for Cambodian students)	No	No
Mclean, D. (2007)	Rangatahi Māori Mentoring Programme	No	No
McInerny, J. (2005)	The Buddy Programme	Yes	Yes
Milne, B. et al., (2002)	Community Intervention Project	No	No
Ministry of Education (2009)	He Ara Tika Māori Mentoring Programme	Yes	Yes
Qiao, C., & McNaught, H. (2007)	Project K	Yes	Yes
Selwood, J. (2005)	Deaf Mentoring Service	No	No
Starpath (2005)	MATES	Yes	Yes
Starpath (2007)	MATES	Yes	Yes
Stevenson, L. (2008)	On-line mentoring of physical education students	No	No
Tasi, B. S. (2009)	Work Transition Mentoring Programme	No	No
Wilson, S. (2006).	Programme to improve retention and success of polytechnic students	No	No
Youth at Risk of Offending Team (2001)	Police Youth at Risk of Offending Programme: One-to-One Mentoring Programme	No	No

Table 4. Mentee and mentor characteristics for included studies.

Citations	Mentee characteristics	Mentor characteristics
Adams, R. J. (2004)	Programme aimed at youth at risk for truancy. Four mentees, male (75%) and female, with a mean age of 11.5 years. Ethnicity of mentees and the duration of the programme were unknown.	The mentors were university student volunteers who were screened and trained and received ongoing monthly support but no supervision of the match.
Afeaki-Mafile'o, E. (2007)	Programme aimed at high risk Pasifika youth. Mentee was one Pasifika female, age unknown. The duration of the programme was 20 months, on average.	The mentors were adult and peer volunteers. It was unknown whether they were screened, trained, or received ongoing match support and supervision.
Ave, K. et al., (1999)*	The mentoring for children/youth at-risk demonstration project. Ethnicity: Maori-27%, Pakeha/NZ Euro-73%.	
	Man Alive (Auckland): Aimed at young males between the ages of 5 and 18 years of age who do not live with their natural father. The expectation for length of relationship was 12 months.	Mentors were adult volunteers who were screened and trained. It was unknown whether they received ongoing support or supervision of the match.
	Just-Us-Youth (Christchurch): Aimed at children whose parent has been incarcerated in a Christchurch prison. An estimated 80-90% of children targeted are of Maori descent. The expectation for length of relationship was 12 months.	Mentors were adult volunteers who were screened and trained. It was unknown whether they received ongoing support or supervision of the match.
	Te Whare Putea (Kaikoura): The expectation for length of relationship was 12 months.	Mentors were adult volunteers who were screened and trained. It was unknown whether they received ongoing support or supervision of the match.
	Presbyterian Support Services (North Otago): Aimed at at-risk children and youth. The expectation for length of relationship was 12 months.	Mentors were adult volunteers who were screened and trained. It was unknown whether they received ongoing support or supervision of the match.
	Tuakana/Teina Mentoring Programme INC (Tararua) Project: The targeted age group was 8-15 years of age (male and female) who was at risk of offending. The expectation for length of relationship is 12 months.	Mentors were adult volunteers who were screened and trained. There was supervision of the match initially but it was unknown whether the mentors received ongoing support.
	Te Runanga O Ngati Porou (East Coast): Aimed at at-risk youth of the Ngati Porou community. The expectation for length of relationship was 12 months.	Mentors were adult volunteers who were screened and trained. It was unknown whether they received ongoing support or supervision of the match.
Ballinger, B. et al., (2009)	Programme for low risk and at-risk young women who showed leadership potential. All female participants (n = 47) were from low-decile schools in the Auckland region, with 15% Pakeha/NZ Euro and 85% of unknown ethnicity. The expectation for length of relationship was 48 months.	The mentors were adult volunteers. All mentors were screened and trained and received monthly ongoing support and supervision of match.
Deane, K., & Harre, N. (2008)	Project K programme for youth with low self-efficacy. Report analyzed 41 Project K programmes across 8 regions. There were 2	The mentors were adult volunteers, mentors were screened and trained and received monthly ongoing

	groups of participants with mixed gender and ethnicity (proportions unknown): Project K group (n = ranged from 263 to 398 depending on the variable measured), Control group (n = 158 to 346). The expectation for length of relationship was 12 months.	support and supervision of match.
Enkey, R. F. (2001)	Project K programme for youth with low self-efficacy. Report analyzed 137 (68 females, 69 males) students from 4 grammar schools under the North Shore Education Trust Project K, with 80% Pakeha/NZ Euro, 9% Māori, 6% Pasifika, and 5% other ethnicity. The expectation for length of relationship was 12 months.	The mentors were adult volunteers, mentors were screened and trained and received monthly ongoing support and supervision of match.

Hammond, J. (2007)	Programme aimed to increase school engagement of students. Eighteen mentees participated, all female, with 28% Māori and 72% Pakeha/NZ Euro. The expectation for length of relationship was unknown.	The mentors were paid adults, mentors were screened and trained. It was unknown whether they received ongoing support and supervision of the match.
Heke, J (2005)	Programme aimed at at-risk Maori youth. Mentees were 25 Māori males (40%) and females between the ages of 11 and 12 years. The expectation for length of relationship was 2.5 months.	The mentors were volunteer peers, mentors were screened and trained and received weekly ongoing support and supervision of match.
Hill, J. (2008)	"I Have a Dream" programme aimed at at-risk low-decile youth. Mentees were 53 males and females between the ages of 12 and 14 year. Mentees' ethnicity was predominantly Māori and Pasifika, proportions unknown. The expectation for length of relationship was 12 months.	The mentors were adult volunteers, mentors were screened and trained and received monthly ongoing support and supervision of match.
Irving, E. et al., (2003)	Programme aimed at high ability Year 13 students. Mentees were 62 students whose gender and ethnicity was unknown. The expectation for length of relationship was 9 months.	The mentors were teachers at the school and were not screened or trained. It was unknown whether they received ongoing support and supervision of the match.
Kostuk-Warren, J. (2005)	Project K programme for youth with low self-efficacy. Twenty-seven mentees (48% male) and a control group (n = 15) made up of Pakeha/NZ Euro (74%), Māori (15%), and other ethnicity (11%) students. The expectation for the length of the relationship was for 12 months.	The mentors were adult volunteers, mentors were screened and trained and received monthly ongoing support and supervision of the match.
Lennan, M. (2006)	Programme aimed at boys of single mothers. Mentees were 11 at-risk males aged between 8 and 18 years. Mentees' ethnicity was predominantly Māori and Pasifika, proportions unknown. The expectation for the length of the relationship was for 12 months.	The mentors were adult volunteers, mentors were screened and trained and received monthly ongoing support and supervision of the match.
Litchfield, J. E. (2006)	Programme aimed at at-risk youth of a mid-decile all boys' school. Mentees were 15 year 9 boys (age 9-14 year) of whom 33% were Maori; 47% were Pakeha/NZ Euro; and 20% were of other ethnicity. The expectation for the length of relationship was 8 months.	The mentors were peer volunteers, mentors were screened and trained and received ongoing support and supervision on the match every 2-4 months.
Lyon, D. R. (1992)	Programme aimed at Cambodian students (n = 13). Age of mentees ranged from 14-20 years of age, with 8 males, 5 females. The expectation for the length of relationship was 2 months.	The mentors were all peer volunteers and all were screened. The mentors were trained and received weekly support but there was no ongoing supervision of the match.
Mclean, D. (2007)	Programme aimed at Maori youth seeking to pursue a career in the health profession. Of the 4 mentees, 3 were Maori and 1 identified as Maori/Pasifika. Mentees were male (50%) and female and aged between 17 and 19 years. The expectation for the length of	The mentors were paid adults, mentors were screened and trained. It was unknown whether they received ongoing support and supervision of the match.

	relationship was 6 months.	
McInerny, J. (2005)	Programme aimed at-risk children between the ages of 4-12 years. The 16 mentees were male (44%) and female and either Pakeha/NZ European (81%) or Maori (19%). The expectation for the length of relationship was 12 months.	The mentors were adult volunteers, mentors were screened and trained and received monthly ongoing support and supervision of the match.
Milne, B. et al., (2002)	Programme for high-risk youth with a history of truancy. The 66 mentees were male (53%) and female with a mean age of 14.8, Pakeha/NZ Euro (65%), Maori (27%), and other ethnicities (8%). The duration of the programme was unclear.	The mentors were adult volunteers, it was unknown whether they were screened or trained or whether they received ongoing support and supervision of the match.
Ministry of Education (2009)	Programme aimed at low achieving Maori youth. The 1074 mentees were male (43%) and female, and typically Maori (98%), age unknown. The expected the length of relationship was 12 months.	The mentors were adult volunteers and were trained and screened, and received monthly ongoing support or supervision of match.
Qiao, C., & McNaught, H. (2007)	Project K programme aimed at youth with low efficacy. The 94 mentees were male (49%) and female between the ages of 14 and 16 years. Ethnicity primarily Pakeha/NZ Euro (53%), Maori (8%), and other (19%). The expected the length of relationship was 12 months.	The mentors were adult volunteers; mentors were screened, trained and received monthly ongoing support and supervision of the match.
Selwood, J. (2005)	Deaf mentoring programme, the 8 mentees age ranged from 6 to 16 years, were of mixed gender (proportions unknown). Mentees' ethnicity and expectation for the length of relationship is unknown.	The mentors were adult volunteers. It was unknown if mentors were screened, trained, and received ongoing support or supervision of the match.
Starpath (2006)	Programme aimed at at-risk youth with the potential to succeed academically. The 96 mentees were male (50.5%) and female and aged between 16 and 18. Their ethnicity was Maori = 22%, Pakeha/NZ Euro (27%), Pasifika (40%), and other (11%). The expectation for the length of relationship was 8 months.	The mentors were paid university students, mentors were screened and trained and received weekly ongoing support and supervision of the match.
Starpath (2007)	Programme aimed at at-risk youth with the potential to succeed academically. The 77 mentees were male and female and aged between 16 and 18. Mentees' ethnicity was predominantly Māori and Pasifika, proportions unknown. The expected relationship length was 8 months.	The mentors were paid university students, mentors were screened and trained and received weekly ongoing support and supervision of the match.
Stevenson, L. (2008)	Online programme for physical education students. Mentees (n = 12) were male (92%) and female students from a Year 13 class. Mentees' ethnicity was 8% Māori and 92% of unknown ethnicity. The expectation for the length of relationship was 2 months.	The mentors were adult volunteers who were experts in their sporting field. They were not screened or trained for the programme and received support and supervision of the match infrequently.

Tasi, B. S. (2009)	Programme aimed to assist at-risk Pasifika youth with the school to work transition. Mentees were 8 Pasifika males with a mean age of 19.6 years. The expectation for the length of relationship was unknown.	The mentors were adult volunteers, it is unknown whether they were screened or trained or whether they received ongoing support or supervision of the match.
Wilson, S. (2006).	Programme aimed to improve retention and graduate rates at New Zealand polytechnic. Mentees were 83 mixed0gender students enrolled at the polytechnic, age and ethnicity were unknown. The expected length of relationship was 4-9 months.	The mentors were staff at the polytechnic. Mentor screening, training and ongoing support and supervision of the match is unknown.
Youth at Risk of Offending Team (2001)	Programme aimed at young offenders and youth at-risk of offending. Mentees were 14 male (57%) and female youth, with a mean age of 12.5years. Mentees were Pakeha/NZ Euro (71%) and Maori (29%). The expectation for the length of relationship was unknown.	Mentors were adults, but it is unknown whether they were paid or volunteers. Mentors were trained. It is unknown whether they were screened. Ongoing support was given monthly while ongoing supervision was given every 2 -4 months.

**Note.* This report incorporated evaluations from multiple studies. Details are provided where information could be extracted on individual programmes.

Table 5. Programme characteristics for included studies.

Citations	**Programme characteristics**							
	Programme dissemination	***Mentoring only vs. component study***	***Site of programme***	***Programme goals***	***Programme type***	***Compensation for mentors***	***Expected frequency of contact***	***Duration of mentoring relationship (minimum for some)***
Adams, R. J. (2004)	Independent	Mentoring only	Community	Academic Behavioural Vocational	One-to-one	Volunteer	Unknown	Unknown
Afeaki-Mafile'o, E. (2007)	Regional multi-site (Affirming Works)	Component	Schools	Academic Vocational	Mixed (one-to-one and group)	Volunteer	Weekly	20 months
Ave, K. et al., (1999)	Man Alive (Auckland)	Unknown	Mixed	Psychological Behavioural Interpersonal	One-to-one	Volunteer	Unknown	12 months
	Just-Us-Youth (Christchurch)	Unknown	Mixed	Psychological Interpersonal	One-to-one	Volunteer	Unknown	12 months
	Te Whare Putea (Kaikoura)	Unknown	Mixed	Interpersonal Vocational	One-to-one	Volunteer	Unknown	12 months
	Presbyterian Support Services (North Otago)	Unknown	Mixed	Academic Psychological Behavioural Interpersonal	One-to-one	Volunteer	Unknown	12 months
	Tuakana/Taina Mentoring Programme INC (Tararua)	Unknown	Mixed	Behavioural Vocational	One-to-one	Volunteer	Unknown	12 months
	Te Runanga O Ngati Porou (East Coast)	Unknown	Mixed	Academic Behavioural Interpersonal	One-to-one	Volunteer	Unknown	12 months
Ballinger, B. et al., (2009)	Regional multi-site (YWCA Future	Component	Mixed (school and community)	Academic Psychological Interpersonal	One-to-one	Volunteer	Fortnightly	48 months

	Leaders)							
Deane, K., & Harre, N. (2008)	National (Project K)	Component	Schools	Academic Psychological Behavioural Interpersonal Vocational	One-to-one	Volunteer	Weekly	12 months
Enkey, R. F. (2001)	National (Project K)	Component	Schools	Academic Psychological Behavioural Interpersonal Vocational	One-to-one	Volunteer	Weekly	12 months
Hammond, J. (2007)	Independent	Component	School	Academic Behavioural	Group	Paid	Weekly	Unknown
Heke, J. (2005)	Independent	Mentoring only	School	Academic Behavioural Vocational Cultural	Group	Volunteer	Weekly	2.5 months
Hill, J. (2008)	Regional multi-site (IHAD Foundation).	Component	Community	Academic	One-to-one	Volunteer	Monthly	12 months
Irving, E. et al., (2003)	Independent	Mentoring Only	School	Academic	One-to-one	Volunteer	Unknown	9 months
Kostuk-Warren, J. (2005)	National	Component	School	Academic Psychological Interpersonal	One-to-one	Volunteer	Monthly	12 months
Lennan, M. (2006)	Regional multi-site (Big Buddy Mentoring Trust)	Mentoring Only	Community	Academic Psychological Behavioural Interpersonal	One-to-one	Volunteer	Monthly	12 months
Litchfield, J.E. (2006)	Independent	Mentoring only	School	Academic Interpersonal	One-to-one	Volunteer	Every 2-4 months	8 months
Lyon, D. R. (1992)	Independent	Mentoring only	Schools	Academic Psychological Interpersonal Cultural	One-to-one	Volunteer	Weekly	2 months
Mclean, D. (2007)	Independent	component	Unknown	Academic Psychological	Group	Paid	Unknown	6 months

				Interpersonal Vocational				
McInerny, J. (2005)	Regional multi-site (The Buddy Programme)	Component	Community	Psychological Interpersonal	One-to-one	Volunteer	Weekly	12 months
Milne, B. et al., (2002)	Independent	Component	Community	Academic Psychological Behavioural Interpersonal	One-to-one	Volunteer	Unknown	Unknown
Ministry of Education (2009)	National	Mentoring Only	Community	Academic	One-to-one	Volunteer	Fortnightly	12 months
Qiao, C., & McNaught, H. (2007)	National (Project K)	Component	School	Academic Psychological Behavioural Interpersonal Vocational	One-to-one	Volunteer	Monthly	12 months
Selwood, J. (2005)	Regional multi-site	Mentoring only	School	Academic Psychological Interpersonal Cultural	One–to-one	Volunteer	Unknown	Unknown
Starpath (2006)	Regional multi-site (MATES evaluation)	Mentoring Only	School	Academic Psychological	Mixed	Paid	Weekly	8 months
Starpath (2007)	Regional multi-site (MATES evaluation)	Mentoring Only	School	Academic Psychological	Mixed	Paid	Weekly	8 months
Stevenson, L. (2008)	Independent	Component	School	Academic Behavioural	Mixed	Volunteer	Less frequent/ never	2 months
Tasi, B. S. (2009)	Independent	Component	Unknown	Vocational	One-to-one	Volunteer	Unknown	Unknown
Wilson, S. (2006).	Independent	Component	School	Academic	One-to-one	Volunteer	Less frequent/never	4-9 months
Youth at Risk of Offending	Independent	Component	Unknown	Academic Psychological Behavioural	One-to-one	Unknown	Monthly	Unknown

Team (2001)				Interpersonal Cultural				

Table 6. Research characteristics for included studies.

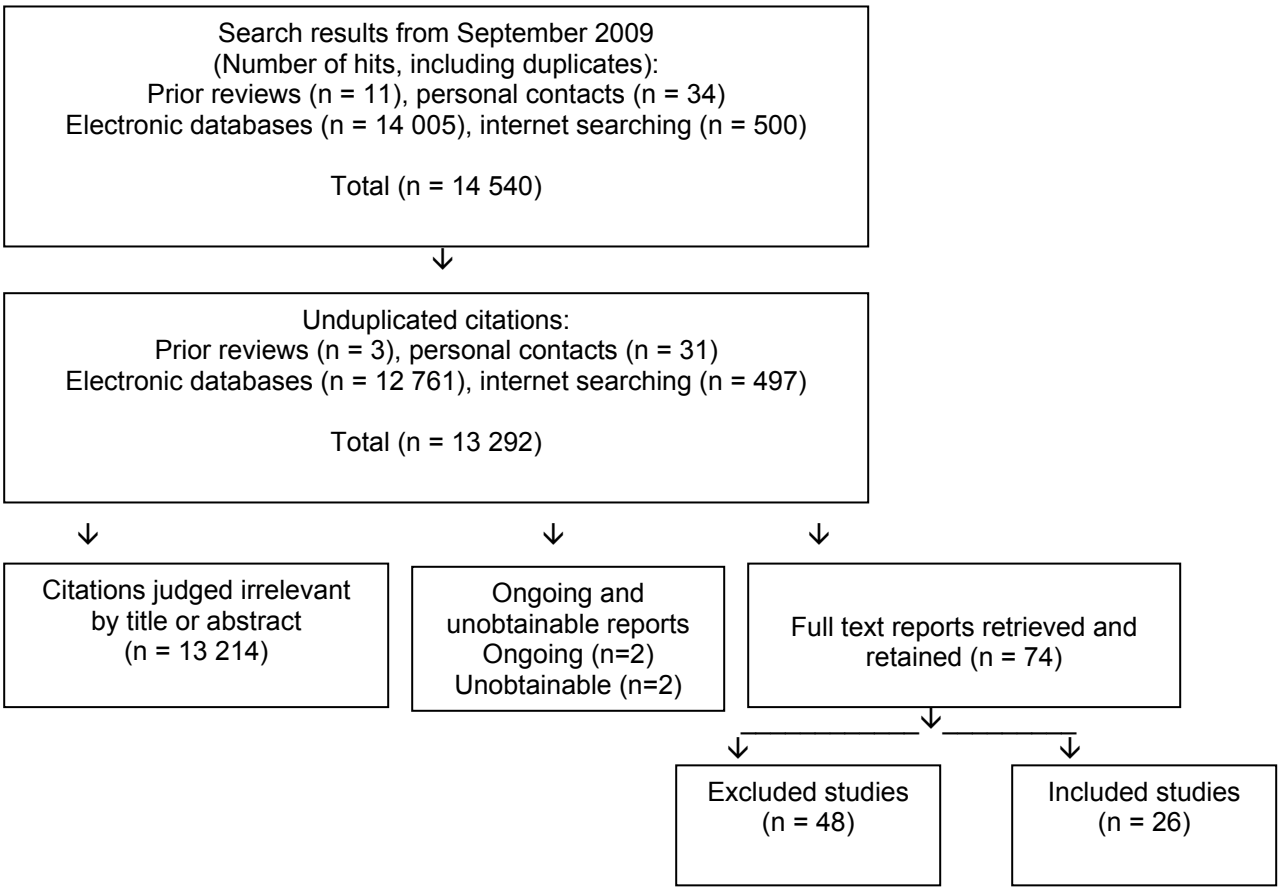
Citations	**Research characteristics**					
	Peer-reviewed	***Type of Publication***	***Author(s)/ Evaluator(s)***	***Methodology***	***Use of a control group***	***Pre-test/ post-test design***
Adams, R. J. (2004)	No	Masters Thesis	External	Qualitative	No	No
Afeaki-Mafile'o, E. (2007)	No	Book chapter	Internal	Qualitative	No	No
Ave, K. et al., (1999)	No	Technical report	External	Qualitative	No	No
Ballinger, B. et al., (2009)	No	Technical report	External	Mixed methods	No	No
Deane, K., & Harre, N. (2008)	No	Technical report	External	Quantitative	Yes	Yes
Enkey, R. F. (2001)	No	Technical report	External	Quantitative	Yes	Yes
Hammond, J. (2007)	No	Postgraduate Diploma research project	Internal	Mixed methods	No	No
Heke, J (2005)	No	Masters Thesis	Internal	Mixed methods	Yes	Yes
Hill, J. (2008)	No	Technical report	External	Mixed methods	Yes	No
Irving, E. et al., (2003)	Yes	Journal article	External	Quantitative	Yes	No
Kostuk-Warren, J. (2005)	No	Doctoral Dissertation	External	Mixed methods	Yes	Yes
Lennan, M. (2006)	No	Technical report	External	Qualitative	No	No
Litchfield, J. E. (2006)	No	Masters Thesis	Internal	Mixed methods	No	No
Lyon, D. R. (1992)	No	Masters Thesis	Internal	Mixed methods	No	Yes
Mclean, D. (2007)	No	Masters Thesis	External	Qualitative	No	No
McInerny, J. (2005)	No	Masters Thesis	Internal	Qualitative	No	No
Milne, B. et al., (2002)	Yes	Journal article	External	Quantitative	No	Yes
Ministry of Education (2009)	No	Technical report	External	Quantitative	No	No
Qiao, C. & McNaught, H. (2007)	No	Technical report	External	Quantitative	Yes	Yes
Selwood, J. (2005)	No	Masters thesis	External	Quantitative	No	No
Starpath (2005)	No	Technical report	External	Mixed methods	No	No
Starpath (2007)	No	Technical report	External	Mixed methods	No	No
Stevenson, L. (2008)	No	Masters thesis	External	Qualitative	No	No
Tasi, B. S. (2009)	No	Masters Thesis	External	Qualitative	No	No
Wilson, S. (2006).	Yes	Journal article	Internal	Quantitative	Yes	No
Youth at Risk of Offending Team (2001)	No	Technical report	External	Mixed methods	No	Yes

Table 7. Number of studies with bias and percentage of bias for included studies.

For all studies (n=26)	**Number of studies with bias**	**Percentage of studies with bias**
Programme participation bias – all eligible participants are offered a place in the programme and at least 80% of those eligible enrol in the programme (n=12).	3	25
Participation bias - at least 80% of those eligible and selected to participate in the research did participate (n=20).	7	35
Avoidance of performance bias - intervention group did not receive any other services beyond mentoring (n=26).	15	58
Avoidance of performance bias - control group did not have any **negative** experiences/services (n=6).	0	0
Attrition bias – programme drop out rate was 20% or less and equal with comparison group if applicable (n=16).	8	50
Conflict of interest - researchers or data collectors would benefit from favourable results of programme (n=26).	7	27
For qualitative studies (n=18)	**Number of studies with bias**	**Percentage of studies with bias**
Selection bias - inclusion of participants (assignment to group) (n=14).	5	36
Detection bias - selected because of favourable results (n=17).	6	35
Intention to treat - includes everyone who started the programme as possible participants (e.g., interviews those who do not complete) (n=18).	10	56
Data analysis techniques are thorough and systematic (n=18).	9	50
Data analysis techniques are guided by a clear, theoretical framework (n=18).	10	56
For quantitative studies (n=18)	**Number of studies with bias**	**Percentage of studies with bias**
Selection bias - inclusion of participants (assignment to group) (n=18).	15	83
Detection bias - assessor unaware of the assigned treatment when collecting outcome measures (n=14).	11	79
Intention to treat- includes those who dropped out of the programme in the analyses (n=15).	11	73
Standardised observation periods – follow-up data were collected from each case at a fixed point in time (after assignment if start point varied) (n=17).	3	18
Used established measures with demonstrated reliability and validity (n=17).	4	24

Note. Due to the large amount of missing data for some categories, the number of studies included in for that particular category is identified in ( ).

Figure 1. Flow chart reflecting search process, screening and inclusion decisions.



APPENDICES

Appendix A: Coding sheets

INCLUSION CHECKLIST

Coder Initials		Date Coded
Study ID	Study Name	Study Authors and Year

	YES	NO
Involves formal mentoring programme, including one-to-one, group, team, peer or e-mentoring.		
Does not focus on informal or natural mentors.		
For qualitative studies, has an indicator of effectiveness reflecting change; post-test only is acceptable if change is discussed.		
For quantitative studies, has an indicator of effectiveness including element of change or difference (e.g., pre-test, post-test change or the use of a comparison group; post-test only studies are excluded).		
For quantitative studies, not a post-test only design without a comparison group (post-test only design may be used, however, if there is an accompanying qualitative section).		
Examines the effectiveness of programme to address outcomes in the following areas: emotional/psychological, problem/high-risk behaviour, academic/educational, career/ employment, and social competence.		
Participants are over the age of 6 years and under the age of 24 years. The mean age for the study is not over 19 years.		
Study is set in New Zealand.		

BIAS ASSESSMENT FOR INCLUDED STUDIES

Coder Initials	Date Coded	
Study ID	Study Name	Study Authors and Year

	MET	UNMET	UNCLEAR
For all studies			
Programme participation bias – all eligible participants are offered a place in the programme and at least 80% of those eligible enrol in the programme	Yes	No	Unclear
Participation bias - at least 80% of those eligible and selected to participate in the research did participate	Yes	No	Unclear
Avoidance of performance bias - intervention group did not receive any other services beyond mentoring *('Yes' if mentoring is embedded in other services; 'No' if programme comprises of mentoring only; 'Unclear' if there is a hint that there may have been other services available to mentees)*	No other services	Yes, other services	Unclear
Avoidance of performance bias - control group did not have any **negative** experiences/services *('Yes' if the mentoring programme is school based and taking place during school hours, or if it is explicitly stated that they are receiving some negative experience/s; 'No' if the programme is not school based and taking place during school hours, and it is not explicitly stated that they are receiving some negative experience/s; 'Unclear' if there is a hint they may be receiving some negative experience/s)*	No other experiences or services	Yes, other experiences or services	Unclear
Attrition bias – programme drop out rate was 20% or less and equal with comparison group if applicable	Yes	No	Unclear
Conflict of interest - researchers or data collectors would benefit from favourable results of programme	No	Yes	Unclear
For qualitative studies			
Selection bias - inclusion of participants (group assignment) *('Targeted selection' yes if only those successful or continuing/ completed the programme are included in the research)*	Random selection/ All selected	Targeted selection	Unclear
Detection bias - selected due to favourable results	No	Yes	Unclear
Intention to treat - includes everyone who started the programme as possible participants (e.g., interviews those who do not complete)	Yes	No	Unclear
Data analysis techniques are thorough and systematic *('No' and 'Unclear' are the same thing)*	Yes	No/Unclear	
Data analysis techniques are guided by a clear, theoretical framework *('Yes' if the framework is explicitly stated; 'No' if no framework is stated ('No' and 'Unclear' are the same thing))*	Yes	No/Unclear	
For quantitative studies			
Selection bias – inclusion of participants (assignment to group)	Random assignment	Matched control	Non-matched control / No control
Detection bias – assessor unaware of the assigned treatment when collecting outcome measures *('No' if there is no control group)*	Yes	No	Unclear
Intention to treat – includes those who dropped out of the programme in the analysis	Yes	No	Unclear
Standardised observation periods – follow-up data were collected from each case at a fixed point in time	Yes	No	Unclear
Used established measures with demonstrated reliability and validity	Yes for all outcomes	Yes for some outcomes / No for all outcomes	Unclear

CULTURAL APPROPRIATENESS CODING SHEET

Coder Initials		Date coded			
Study ID	Study Name		Study Authors and Year		
Ethnicity	% Māori	% Pakeha/NZ Euro	% Pasifika	% Other	Unknown
Programme developed for a particular ethnic group(s)		Yes	No	If yes, which group(s):	

Cultural appropriateness					*Notes*
Researcher(s)	Appropriate	Somewhat appropriate	Not appropriate	Ignored	
Research procedures and analysis	Appropriate	Somewhat appropriate	Not appropriate	Ignored	
Programme implementers	Appropriate	Somewhat appropriate	Not appropriate	Ignored	
Programme design	Appropriate	Somewhat appropriate	Not appropriate	Ignored	

Appropriate: acknowledgement of cultural issues and demonstration that they were taken into consideration; provided information about the researcher's and implementer's cultural backgrounds; there was cultural competency training of the people delivering the programme; analysis took into account the cultural background of the participant
Somewhat appropriate: acknowledgement of cultural issues but no significant demonstration that they were taken into consideration
Inappropriate: acknowledgement of cultural issues but incorrect response to these issues
Ignored: no acknowledgement of cultural issues or attempt to take them into consideration by programme implementers or researchers, no information provided about researchers or implementers

DATA EXTRACTION CODING SHEET

Coder Initials		Date coded									
Study ID	Study Name								Study Authors and Year		
Notes/ Page #											
Report information											
	Year of study				Unknown						
	Peer-reviewed publication *('No' if it is a thesis, dissertation or private/technical report)*	Yes			No				Unclear/Unknown		
	Author(s)/Evaluator(s)	External			Mixed				Internal		
	Type of publication	Journal article	Book chapter		Thesis/ dissertation	Technical report			Conference paper	Other	
	Role of evaluator/author in the programme	Evaluator delivered therapy/ treatment	Evaluator designed programme or was involved in planning, controlling or supervising delivery of programme		Evaluator influential in service setting but no direct role in delivering, controlling or supervising programme			Evaluator independent of service setting and treatment – research role only		Cannot tell	
Methodology											
	Sample size at start – treatment				Unknown						
	Sample size at start – control				Unknown	NA					
	Attrition rate – treatment				Unknown						
	Attrition rate – control				Unknown	NA					
	Design type *('Experimental' = 2 groups, random assignment, pre-test post-test design, manipulation of the key variable; 'Quasi-experimental with control' = 2 groups, matched or non-matched; 'Quasi-experimental without control' = one group, pre-test post-test design; 'Correlational' = one group, post-test only)*	Experimental	Quasi-Experimental with control		Quasi-Experimental without control			Correlational		Qualitative	
	Research methodology	Quantitative			Qualitative				Mixed		
Programme features											
	Programme dissemination *('International' = programme is found internationally, e.g., Big Brothers; 'National' = programme is found nationally, e.g. Project K; 'Regional multi-site' = programme only in a regional area, e.g., MATES; 'Independent' = programme developed as a one-off for a specific site)*	International	National		Regional multi-site			Independent			

	Age of programme	Relatively new (e.g., less than 2 years old, or few client cohorts)	Established programme (e.g., established more than 2 years, or many client cohorts)	Defunct programme, evaluated post hoc	Cannot tell
	History of programme evaluation (internal or external) *('Yes' if the programme has had consistent or previous internal or external evaluations carried out, e.g., 6 monthly mentor and mentee surveys, evious evaluation)*	Yes	No		Unclear/Unknown
	Programme is based on principles of good practice *('Yes' if based on a proven established programme, theory/frameworks influenced formation of programme that are research-based)*	Yes	Somewhat	No	Unclear/Unknown
	Stand-alone versus component	Stand-alone	Component		Unclear/Unknown
	For component studies only, proportion of all services that was mentoring	Most	About half	Less than half	Unclear / NA
	Programme goals are identified	Yes	No		Unclear
	Programme general location	Urban	Rural	Mixed	Unclear
	Programme site	School	Community centre/ private facility	Juvenile justice facility	Mixed / Unclear
	Education component to programme *e.g., tutoring, monitoring attendance, homework help, special interest classes*	Yes	No	Unclear	NA
	Counselling component to programme *e.g., case management, grief groups, psychological services and counselling)*	Yes	No	Unclear	NA
	Interpersonal/personal skills component to programme *e.g., leadership skills, team building*	Yes	No	Unclear	NA
	Cognitive skills component to programme *e.g., moral reasoning*	Yes	No	Unclear	NA
	Behavioural component to programme *e.g., rewards for positive behaviour, physical activity*	Yes	No	Unclear	NA
	Employment component to programme *e.g., career counselling, job training, non-paid work service, job placement supervision*	Yes	No	Unclear	NA
	Life skills/needs component to programme *e.g., voluntary community service work, goal setting, outward bound and camps*	Yes	No	Unclear	NA

	Cultural component *e.g., addresses issues of a particular cultural group*	Yes	No	Unclear	NA	
	Programme type	One-to-one	Group	Mixed		
	Relationship type	Peer	University Student	Other Adult	Other	Unknown
	Compensation for mentors *(assume volunteer if they are peer-mentors)*	Paid	Non-monetary compensation	Volunteer	Unknown	
	Mentors are screened	Yes	No	Unclear		
	Mentors are trained	Yes	No	Unclear		
	Ongoing support for mentors	Weekly	Monthly	Every 2-4 months	Less frequent or never	Unknown
	Ongoing supervision of match	Weekly	Monthly	Every 2-4 months	Less frequent or never	Unknown
	Established criteria for matching mentors and mentees	Yes	No	Unclear	NA	
	Level of structure	Highly structured *(Location, timing, activities)*	Moderate *(Location/ time, not activities)*	Minimal *(Time expectation only)*	Little/No structure	Unclear
	Inclusion of families/parents/caregivers	Very Frequent (Weekly)	Frequent (Monthly)	Moderate (Every three months)	Minimal (Once or twice per year)	No contact	Unclear
	Expectation for frequency of contact	Weekly	Every other week	Monthly	Every other month	Less than every other month	Unclear
	Expectation for duration of each contact (minutes)		Unclear/Unknown			
	Expectation for length of relationship (months)		Unclear/Unknown			
Youth characteristics						
	Gender	Male Only	Female only	Mixed, % Male	Unknown	
	Ethnicity	% Maori	% Pakeha/NZ Euro	% Pasifika	% Other	Unknown
	Age of participants					
	Mean Age		Unknown			
	Age range		Unknown			
	Socioeconomic background	High (Decile 8-10)	Mid (Decile 4-7)	Low (Decile 1-3)	Mix	Unknown
	Risk status	Typical/ community, low/no risk	At-risk: poverty, school problems, family problems, low self-esteem, without severe problems	High-risk: youth offenders, substance problems, clinical/ mental health problems, severe family problems, educational failure, institutional setting	Mixed	Unknown
Mentor-Mentee relationship						
	Average length of relationship (months)		Unclear			

	Average frequency of contact, face to face	Weekly	Every other week	Monthly	Every other month	Less than every other month	Unclear
	Average length of each contact, face to face (minutes)			Unclear/Unknown			
	Average frequency of contact, non-face to face	Weekly	Every other week	Monthly	Every other month	Less than every other month	None	Unknown
	Types of non-face to face contact	Phone/text	Social Networking	Email	NA		
	Matched on gender	Yes	No	Unclear/ Unknown	NA		
	Matched on ethnicity	Yes	No	Unclear/ Unknown	NA		
Outcome goals and measures		**Goals**			**Measures**		
	Academic improvement and school adjustment *e.g., achievement, grades, attendance, graduation, attitudes*	Yes	No	Unclear	Yes	No	Unclear
	Psychological adjustment – internalising problems *e.g., self-esteem, confidence, depression, help-seeking*	Yes	No	Unclear	Yes	No	Unclear
	Behavioural adjustment – externalising problems *e.g., offending rates, physical activity, substance use*	Yes	No	Unclear	Yes	No	Unclear
	Interpersonal adjustment *e.g., social self-efficacy, family cohesion, leadership skills*	Yes	No	Unclear	Yes	No	Unclear
	Vocational adjustment *e.g., career decision self-efficacy, employment status,*	Yes	No	Unclear	Yes	No	Unclear
	Cultural adjustment *e.g., cultural identity*	Yes	No	Unclear	Yes	No	Unclear
	Source of information *Circle all that apply*	Self-report	Third party report	Archival data/records	Observation		
	Outcome measures cover programme goals	Yes, All	Some, more than half	Some, less than half	None	No programme goals stated	Unclear
Adverse outcomes							
	Occurrence of adverse outcomes	Yes	No	Unclear/ unknown			
	Types of adverse outcomes						
	Academic/school	Yes	No	NA			
	Psychological	Yes	No	NA			
	Behavioural	Yes	No	NA			
	Interpersonal	Yes	No	NA			
	Vocational	Yes	No	NA			
	Cultural	Yes	No	NA			

Timing of intervention assessment							
	Quantitative						
	Pre-test	Yes	No	Unclear/ Unknown	NA		
	Mid-intervention	Yes	No	Unclear/ Unknown	NA		
	Immediate post	Yes	No	Unclear/ Unknown	NA		
	Short term follow up (2-8 weeks)	Yes	No	Unclear/ Unknown	NA		
	Moderate follow up (2-6 months)	Yes	No	Unclear/ Unknown	NA		
	Long-term follow up (6+ to 12 months)	Yes	No	Unclear/ Unknown	NA		
	Long-term follow up (12+ months)	Yes	No	Unclear/ Unknown	NA		
	Qualitative						
	Pre-test	Yes	No	Unclear/ Unknown	NA		
	Mid-intervention	Yes	No	Unclear/ Unknown	NA		
	Immediate post	Yes	No	Unclear/ Unknown	NA		
	Short term follow up (2-8 weeks)	Yes	No	Unclear/ Unknown	NA		
	Moderate follow up (2-6 months)	Yes	No	Unclear/ Unknown	NA		
	Long-term follow up (6+ to 12 months)	Yes	No	Unclear/ Unknown	NA		
	Long-term follow up (12+ months)	Yes	No	Unclear/ Unknown	NA		
Quantitative studies only							
	Outcomes						
	Academic/school adjustment effect (1)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Academic/school adjustment effect (2)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Academic/school adjustment effect (3)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Academic/school adjustment effect (4)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Psychological adjustment effect (1)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Psychological adjustment effect (2)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation

	Psychological adjustment effect (3)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Psychological adjustment effect (4)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Behavioural adjustment effect (1)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Behavioural adjustment effect (2)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Behavioural adjustment effect (3)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Behavioural adjustment effect (4)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Interpersonal adjustment effect (1)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Interpersonal adjustment effect (2)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Interpersonal adjustment effect (3)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Interpersonal adjustment effect (4)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Vocational adjustment effect (1)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Vocational adjustment effect (2)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Vocational adjustment effect (3)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Cultural adjustment effect (1)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Cultural adjustment effect (2)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation
	Cultural adjustment effect (3)	Outcome	Statistics	Effect size	Proportion success	NA	Explanation of calculation

	Control						
	For studies with control groups	Random assignment	Matched assignment – matched on outcomes	Matched assignment – matched on demographics	No matching	NA	
	What's done to the control group	Nothing	Waitlist, delayed treatment control	Other services as usual	Other treatment	Unclear	NA
	Limitations to quantitative study						
Qualitative studies only							
	Research design						
	Type of study *Circle all that apply*	Single case-study	Multiple case-study	Interviews	Focus group	Unclear	NA
	Research paradigm/epistemology clearly stated	Yes	No/Not stated	NA			
	Research matches stated epistemology *i.e., regarding overall design – how participants are selected, interviews conducted, analysis approach*	Yes	No/Not stated	NA			
	Methods used are appropriate for the research question	Yes	No/Not stated	NA			
	Design developed and adapted to social context as needed	Yes	No	Not needed	NA		
	Researcher						
	Researcher is sensitive to the social context of the participants	Yes	No/Not stated	NA			
	Reflexivity acknowledged *(researcher acknowledges their individual characteristics and role as a researcher, and how this impacts interactions with participants and conclusions drawn about data)*	Yes	No/Not stated	NA			
	Sampling						
	Subsample	Yes	No	Unclear/not stated	NA		
	Participants clearly defined and identified	Yes	No	Unclear	NA		
	Sampling strategies appropriate for intended sample	Yes	No	Unclear	NA		
	Sample adequately reflects intended target group	Yes	No	Unclear	NA		

	Data collection and analysis							
	Methods and procedures clearly described	Yes		No		NA		
	Data analysis process is clearly described	Yes		No		NA		
	Data analysis techniques are thorough and systematic	Yes		No		Unclear	NA	
	Data analysis includes voice of the participant through the use of quotes	Yes		No		NA		
	There is a wide variety of views expressed in the analysis	Yes		No		NA		
	Researcher acknowledges contradictions	Yes		No		Unclear	NA	
	Outcomes						***Notes***	
	Academic/school effect	Addressed and effective	Addressed and not effective	Addressed and mixed results	Not addressed	Not a goal of programme		NA
	Psychological effect	Addressed and effective	Addressed and not effective	Addressed and mixed results	Not addressed	Not a goal of programme		NA
	Behavioural effect	Addressed and effective	Addressed and not effective	Addressed and mixed results	Not addressed	Not a goal of programme		NA
	Interpersonal effect	Addressed and effective	Addressed and not effective	Addressed and mixed results	Not addressed	Not a goal of programme		NA
	Vocational effect	Addressed and effective	Addressed and not effective	Addressed and mixed results	Not addressed	Not a goal of programme		NA
	Cultural effect	Addressed and effective	Addressed and not effective	Addressed and mixed results	Not addressed	Not a goal of programme		NA
	Interpretation and presentation of findings							
	Interpretation follows logically from the analysis	Yes		No		NA		
	Findings match the derived data	Yes		No		NA		
	Adequate proportion of data taken into account when presenting findings	Yes		No		Unclear	NA	
	Limitations are noted *(e.g.,generalisibility)*	Yes		No		NA		
	Utility of the findings are noted	Yes		No		NA		
	Limitations to qualitative study							

Appendix B: Review Protocol

Review Protocol–: Mentoring and Young People Systematic Review

Prepared for:

HRC
Ministry of Youth Development

Prepared by:

Susan Farruggia
School of Teaching, Learning and Development
Faculty of Education
The University of Auckland

Table of Contents

1. BACKGROUND FOR THE REVIEW 86
2. OBJECTIVES OF THE REVIEW 87
3. METHODS 88
3.1 CRITERIA FOR INCLUSION AND EXCLUSION OF STUDIES IN THE REVIEW 88
3.2 SEARCH STRATEGY FOR IDENTIFICATION OF RELEVANT STUDIES 89
3.3 DESCRIPTION OF METHODS USED IN COMPONENT STUDIES 89
3.4 CRITERIA FOR DETERMINATION OF INDEPENDENT FINDINGS 90
3.5 DETAILS OF STUDY CODING CATEGORIES 90
3.6 STATISTICAL PROCEDURES AND CONVENTIONS 91
3.7 TREATMENT OF QUALITATIVE RESEARCH 92

4. TIMEFRAME 92
5. PLANS FOR UPDATING THE REVIEW 93
6. ACKNOWLEDGEMENTS 93
7. CONFLICT OF INTEREST 93
8. REFERENCES 94
9. TABLES 94

Background for the Review

Mentoring is "an enduring relationship between a novice and an older, more experienced individual who provides guidance in a particular domain" (Evans & Ave, 2000, p. 41). Evans and Ave (2000) have summarized the youth mentoring literature and have identified five mechanisms through which mentors can promote change within a young person. These include serving as a role model to demonstrate qualities and behaviours for the young person to imitate and internalize; acting as a substitute parent for youth who either don't have a parent or whose relationship with their parents are poor; providing social support, especially emotional support; developing specific, positive skills, such as those needed in the career and work domain; and, finally, modifying undesirable behaviours, such as improving academic achievement and motivation and decreasing involvement in problem behaviour.

Further, five types of mentoring have been identified including: the traditional one-to-one mentoring pair, group mentoring, team mentoring, peer mentoring and e-mentoring (MENTOR/ National Mentoring Partnership, 2005). Traditional one-to-one mentoring involves one adult matched with one child in which the pair typically meets for at least four hours per month for at least a year; some exceptions, such as school-based mentoring programmes, exist in which the duration is often shorter. In group mentoring, one adult mentor is joined with up to four youth to develop mentoring relationships; typically, the mentor serves as a leader for group-based activities. For team mentoring, several adults work with several groups of young people. There is often fluidity in the mentoring units; however, the adult-youth ratio is typically no more than one to four. Peer mentoring occurs when a young person mentors a younger person. Typically, peer mentoring is school-based with an older student mentoring a younger student during school hours. Finally, e-mentoring involves one adult forging a relationship over the internet with one youth; sometimes they have a few face-to-face meetings. Typically, the relationship is focused around school or career support and sometimes can serve as a bridge during the summer holiday period for other types of mentoring.

Researchers have conducted studies to look at the impact of mentoring on at-risk populations. Involvement in youth mentoring programmes has been found to be to associated with less absence from school, more positive attitudes toward school, greater well-being, a more positive reaction to situations involving drugs (LoSciuto, Rajala, Townsend, & Taylor, 1996), less likelihood to start using illegal drugs and alcohol, less engagement in aggressive behaviour, decreasing in wagging school, and lying to their parents (Grossman & Tierney, 1998).

DuBois and colleagues (2002) conducted a meta-analysis examining the effectiveness of one-to-one mentoring programmes in the U.S. Reviewing 55 evaluations of youth mentoring programmes, they identified characteristics of the most effective programmes, including: strong relations between youth and mentors, using mentors from "helping" backgrounds, providing ongoing training and support to the mentors, involving parents, programmes that are based on both theory and research, and targeting at-risk (versus typical) youth. They found that matching on gender, race or interest, as are commonly practiced among many mentoring programmes, did not impact the effectiveness of the programme.

In examining the New Zealand context of mentoring, Evans and Ave (2000) point out that practices of mentoring in the U.S. do not necessarily fit with the New Zealand familial/social structure. While the U.S. programmes typically involve one-on-one relationships, this may not be appropriate for youth in New Zealand where this practice may conflict with social and cultural structures, such as *whanau*. Mentoring for young people in New Zealand needs to account for the cultural needs and practices of its youth. In relation to Maori, acknowledging the importance of personal Maori identity is an issue that has to be addressed, particularly with youth. By helping the young person to know of his/her whakapapa, you are providing the basis of addressing self confidence and awareness; if this can be expressed in te reo, they become much stronger. In addition, how Maori hui, how meetings are run is important. Specifically, providing opportunity to self-identify within the Maori cultural context at the outset of a hui, family involvement, opening with karakia, and creating situations of involvement rather than just being spoken to are essential. In relation to Pasifika youth, they have a need for an

understanding of personal/familial history for their young people. As a migrant community, the young people find themselves balancing the values of western society with the traditional values which are being reinterpreted by their parents. There are also language challenges in that many of the young people do not speak their 'mother tongue' at home as a result of a move in the 1980's by many schools to discourage Pasifika parents from speaking their language at home. Pasifika youths' needs are located in an identity/culture perspective, but also in the migrant nature of their traverse to NZ. These issues have important implications for providers and will be included in the coding scheme for this review, such as type of delivery (traditional versus group), incorporation of cultural practices, and ethnicity of mentors.

Objectives of the Review

This project aims to examine the effectiveness of youth mentoring programmes in New Zealand and identify the characteristics of successful programmes. Further, the study aims to identify gaps in the literature and recommend directions for future research.

Methods

Criteria for inclusion and exclusion of studies in the review

Inclusion and exclusion criteria of studies will be based on a range of categories including: type of programme, participant characteristics, an examination of the effects of the programme, and quality of the study. The Campbell Collaboration systematic review on mentoring by Tolan, Henry, Schoeny, and Bass (2008) will be used to inform and shape our systematic review, particularly in the development of sample spreadsheets and coding criteria. The systematic review will also incorporate information gleaned from this GRADE website, as appropriate, on study design.

Mentoring programmes that will be included in this review are: the traditional one-to-one mentoring pair, group mentoring, team mentoring, peer mentoring and e-mentoring. This inclusion will allow for the diversity seen in youth mentoring programmes in New Zealand. However, it will not include programmes or relationships which fall out of the traditional mentor role, such as coaches or mental health professionals. While we acknowledge the importance of these informal mentors, based on the international literature and as directed by the MYD this review will focus on the effects of *formal* mentoring programmes. Formal mentoring programme will be defined as a programme that facilitates the "process by which a more experienced, trusted guide forms a relationship with a young person who wants a caring, more experienced person in his/her life, so that the young person is supported in growth towards adulthood and the capacity to make positive social connections and build essential skills is increased." (TYMT, 2008).

For the participant characteristics, the primary criterion is age of the youth. Based on international standards (DuBois et al, 2002) and on MYD's definition of youth, studies involving all youth under the age of 24 and over the age of 6 will now be included in the review.

Programme effects will reflect the broad influence that international research has indicated that youth mentoring has. Therefore, studies that examine the effectiveness of emotional/psychological, problem/high-risk behaviour, academic/educational, career/ employment, and social competence (DuBois et al., 2002) will be included. It is not anticipated that any studies will be excluded based on the types of outcomes they are addressing.

The final criterion will be the quality of the study. As this review recognises that there are likely to be major gaps in the New Zealand mentoring literature, a wider array of literature will be included, even if there are flaws with the scientific rigour of the methodology. We plan to include all studies in the review, regardless of design, as the body of literature is quite small. However, limitations of less rigorous studies (e.g., non-experimental or quasi-experimental) will be clearly noted and coded. For quantitative studies, all studies must indicate that some examination of change or difference is included. This is seen as the utilisation or pre-test, post-test change or the use of a comparison group. Please see below for a more detailed description of the treatment of qualitative data. Flaws in the research methodology will be identified in the review and discussed as a limitation.

Search strategy for identification of relevant studies

The research team will follow the guidelines discussed in the Julia Littell workshop. The following procedures will be taken to secure all possible research on youth mentoring in New Zealand:

1. Search of national and international databases. In this search, international literature will be used to inform our framework, however, only New Zealand literature will be included in the review. While a professional librarian will be consulted, this search will include databases such as PsycInfo, Eric and ProQuest Dissertations and Theses. It will use keywords as both subject terms (e.g., mentoring) and text words (e.g., Project K). The following are examples of key terms to be included and refined: *Programmes*: Impact, effect, ineffect, outcome, eviden, success, unsuccess, fail, benefit, achieve, result; *Youth*: Young person, adoles, youth, juvenile, teen, rangatahi, young people, young adult, emerging adult, young, delinquent; *Mentoring*: Mentor, intervention, approach, program, role model; *Outcome*: achieve, self-esteem, depress, social outcome, employ, unemploy, inclus, exclus, engage, disengage, participat, involve, resilien, strength, at risk, offend, displace, vulnerable, problem, misconduct, delinquenAs noted by R:06 research literature in languages other than English will be considered for inclusion. Specific team members will be able to advise on documents in Maori and Samoan. If required, the team will seek support from Faculty of Education colleagues in translating other languages.
2. A general search using Google will also be conducted using subject terms or text words limiting the search with “New Zealand”.
3. All New Zealand university databases will be searched as much research on youth mentoring has been published as theses.
4. Government and university staff webpage’s will also be examined to find mentoring research that has not been published in journals.
5. In addition, The Youth Mentoring Trust (TYMT; see Ann Dunphy under Research Team) will correspond with all mentoring programmes within their umbrella to request evaluation work that is not published or publicly accessible.

Description of methods used in component studies

Mentoring is often only one component of an intervention programme. In terms of evaluation, it is likely the mentoring component will form part of the overall programme evaluation. Therefore, it is important that the review include both stand alone studies and those that form part of a group of wraparound services (i.e., component studies). It is acknowledge that it will be difficult to tease apart effects of mentoring when it is embedded within other wrap-around services. However, it is important to ensure these studies are included as the exclusion would lead to missing many programmes. Consequently, we are developing coding strategies to account for this (see 3.4 following). Locating and evaluating component studies will require special attention as they are likely to be embedded within the context of a larger study and consequently may be more difficult to locate. In addition, it may be difficult to isolate the effect of the mentoring component. This will be noted and taken into account when conducting the review, and a level of independence, in terms of method and analysis, will be necessary to fully evaluate the effectiveness of a component study.

Unlike most international reviews, this review will consider studies using qualitative methods, as this methodology fits will within the cultural traditions of NZ, particularly for Maori and Pacific peoples. As noted in section 3.1 above, it is our intention to include a wide range of quantitative studies at various levels of constraint, from randomised controlled trials to pre- and post-test comparisons without controls.

Criteria for Determination of Independent Findings

Each study will be carefully reviewed to determine the relationship of the researcher(s) to the programme, especially within the context of internal evaluations. This will be important as it will identify the independence of research findings and any potential conflicts of interest that may exist. Special consideration will be applied to qualitative research, especially action research where stakeholders, mentors, and/or mentees are often part of the research team.

To account for different outcome measures, we will categorise into separate domains base on the type of outcome (e.g. school, family). Taking school as an example, this domain will be further categorised and coded into attendance, motivation, academic reports, and whether the outcome was obtained though self- or other-report. In addition, all outcomes will be coded for relevance based on the programme goals, and reliability; and we will also code for measurements that occur at different time points to ensure the findings do not overinflate the effect. Studies will also be coded for sample pool to ensure the same participant is represented only once. If more than one study is conducted using the same participants and the same outcome measure, the more rigorous study will be selected for inclusion in the review. These clarifications will allow us to stratify by study design and quality; in addition, we will take into consideration how this may impact effect size.

Details of study coding categories

Each study or report will be coded on a range of characteristics. These will reflect six major categories that have been adapted from international studies of youth mentoring (e.g., DuBois et al., 2002) and will be guided by The Cochrane Handbook, including:

1. report information (year of study, published/unpublished);
2. evaluation of methodology (qualitative versus quantitative, type of design, sample size, attrition rate, and internal versus external evaluation);
3. programme features (stand-alone versus component of wrap-around, programme goal(s), national programme versus independent programme, urban versus rural programme, compensation of the mentors, degree of monitoring of the programme implementation, mentor characteristics, mentor screening procedures, mentor-mentee matching criteria, mentor training, supervision, support to mentors and mentees, expectations for frequency of contact, duration of contact and length of relationship (including average length of the relationship and expected length of the relationship), level of structure of activities, and inclusion of parents/families);
4. characteristics of youth (gender, ethnicity, mean age, socioeconomic background, at-risk status);
5. mentor-mentee relationship (actual amount of contact, quality of contact);
6. outcome assessment (type of outcome, sources of information on outcomes,We include in our coding a range of outcomes to capture differing programme goals, plus a code for meeting programme goals);
7. adverse outcomes (occurrence of any adverse outcomes); and
8. timing of intervention assessment (whether the assessment occurs during the intervention, an immediate post-test, a short term follow-up (within 6 weeks) and a long term follow-up (more than 6 weeks); and the duration of the long term follow-up: less than one year or more than one year).

In addition to these categories, for quantitative data, any statistics provided that reflect programme effectiveness will also be extracted. This can include changes in mean scores (given with standard deviations), and effect sizes.

The following process will be used to check the accuracy of the coding. Inclusion and exclusion of studies will be determined by two independent coders (SI and RA). Whenever agreement cannot be reached, the PI will make the final decision. For coding of the studies, the process involves a number of steps. The first five included studies will be coded by two independent reviewers (SI and RA). After each one is coded, the SI and RA will compare coding. Any discrepancies in coding will be discussed with the PI, and 3-5 investigators as appropriate. If by study five, agreement is 90% or higher upon initial coding, the consensus process will then be conducted after every 5th study. However, all studies will be double coded, checked for agreement, and discrepancies will be determined by the PI and investigators 3-5 as appropriate.

Statistical procedures and conventions

This review will not discount using meta-analysis; however, it is suspected that the quality and quantity of the available literature may not allow us to conduct a meta-analysis. Should these suspicions prove incorrect, conducting a meta-analysis will be investigated as a future direction as advised. That said, all studies that include statistical evidence of effectiveness will be discussed in terms of the level of effectiveness. Where possible, mean scores will be changed to effect sizes so that uniform comparisons can be made. Where available, we will convert statistics to Cohen's *d* and we will include codes for the effectiveness of programmes. We will use the computer programme RevMan.

Treatment of Qualitative Research

It is anticipated that a significant proportion of research, particularly in New Zealand, will be qualitative in nature (i.e. interview, focus group, or observation). It is therefore important to establish a set of criteria for evaluating this research for inclusion in the review. Studies will be considered across all paradigms – realist or critical. The criteria used to evaluate qualitative research is based on recommendations as noted in the literature (e.g., Elliott, Fischer, & Rennie, 1999; Fossey, Harvey, McDermott, & Davidson, 2002; Lincoln & Guba, 1985). It is acknowledged that the criteria used to evaluate qualitative research is dependent upon the assumed paradigm of the researcher, and will take this into account during the evaluation process. The criteria are as follows:

1. *Research design*: Does the research clearly state the philosophical paradigm/epistemology taken? Does the conducted research match the stated paradigm? Are the methods used appropriate for the research question? Did the design develop and adapt to the social context as needed?
2. *Researcher*: Is the researcher(s) sensitive to the social context of the participants? Is reflexivity acknowledged?
3. *Sampling*: Are the intended participants clearly defined/identified? Are the sampling strategies used appropriate for the intended sample? Does the sample adequately reflect those identified as intended participants?
4. *Data collection and analysis*: Are the methods and procedures used clearly described? Was a systematic process used for the analysis and is this process clearly described? Does the analysis include the voice of the participants through the use of quotes? Are there a range or variety of views expressed? Does the researcher acknowledge contradictions?
5. *Interpretation and Presentation of findings*: Does the interpretation follow logically from the analysis? Do the findings match the derived data? What proportion of the data is taken into account when presenting the findings? Are limitations noted? Are any statements made about utility and generalisability?

Timeframe

The timeframe for the project is as follows:

Task	Responsibility	Date
Write final version of review protocol	Sue	31 Aug 2009
Develop sample spreadsheets	Sue/Pat	31 Aug 2009
Search databases	RA	30 Sept 2009
Request unpublished reports from programmes	Ann/Efeso/Frank	30 Sept 2009
Studies are evaluated for inclusion/exclusion	All	31 Oct 2009
Studies are assessed for bias	All	31 Oct 2009
Data are extracted from studies onto standardised forms	RA	20 Dec 2009
Data are examined and synthesised	All	20 Dec 2009
Characteristics of effective programmes are identified	Sue/Pat	20 Dec 2009
Tables of study information are created	RA	20 Dec 2009
Graphs are made	RA	20 Dec 2009
Discussion of findings	All	20 Dec 2009
Draft report	Sue/Pat/RA	31 Jan 2010
Finalise report	All	20 Feb 2010

Plans for Updating the Review

This review will be developed in a manner that will allow for straightforward updates as the review protocol and final report will clearly indicate what procedures were used and what the findings are. As part of the recommendations for future research, a timeline will be established for updating the review. It is the intention of this research team to update the review, likely in a 3 to 5 year time period (again, based on the timelines suggested by the review).

Acknowledgements

MYD and HRC will be acknowledged in all subsequent outputs (e.g. publications, presentations) from this review.

Conflict of Interest

There are no conflicts of interest.

References

DuBois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: a meta-analytic review. *American Journal of Community Psychology, 30*(2), 157-197.

Elliott, R., Fischer, C. T., & Rennie, D. L. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British Journal of Clinical Psychology, 38*, 215-229.

Evans, I. M., & Ave, K. T. (2000). Mentoring children and youth: Principles, issues, and policy implications for community programmes in New Zealand. *New Zealand Journal of Psychology, 29*(1), 41-49.

Fossey, E., Harvey, C., McDermott, F., & Davidson, L. (2002). Understanding qualitative research. *Australian and New Zealand Journal of Psychiatry, 36*, 717-732.

Grossman, J. B., & Tierney, J. P. (1998). Does mentoring work? An impact study of the Big Brothers Big Sisters program. *Evaluation Review, 22*(3), 403-426.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. California.

LoSciuto, L., Rajala, A. K., Townsend, T. N., & Taylor, A. S. (1996). An outcome evaluation of Across Ages: An intergenerational mentoring approach to drug prevention. *Journal of Adolescent Research, 11*(1), 116-129.

MENTOR/ National Mentoring Partnership (2005). The elements of effective practice. Retrieved from http://www.mentoring.org/program_staff/files/elementseffectivepractice.pdf

Tables

Sample spreadsheets will be provided at a later date.