

This chapter documents the procedure through which the studies that are analyzed in *Chapter Four* were selected.

### *3.1. Identification of applicable studies*

Relevant studies were systematically discovered by electronically searching databases and journal archives and, after applying a series of inclusion and exclusion criteria to the resulting research, using Weight of Evidence (WoE) criteria to determine the level of significance to this review of each paper. The research was available electronically via Oxford University's Oxford Libraries Information Platform (OxLIP+) and did not require disbursement or the expressly solicited permission of the researcher, due to financial constraints. The articles were not necessarily available in the public domain; some papers required a subscription and permissions were obtained through OxLIP+. In order to ensure that all included studies would be equally accessible, only journal articles available in an electronic format were included.

August and Hakuta (1997) claim that the 'cognitive revolution of the 1960s' (2) and the United States' Bilingual Education Act of 1968 led to the first phase of research in the area of second language learner education in the early 1970s, particularly after the 1974 United States Supreme Court judgment in *Lau v. Nichols* which ruled that providing non-English-proficient or limited-English-proficient students with educational materials in English did not constitute 'equality of treatment' ('Lau v. Nichols 26', quoted in Hakuta & Garcia 1989). However, Cummins (1986) notes that, although the development of bilingual programs in the United States began in the 1960s, 'previous attempts at educational reform ha[d] been unsuccessful' (18). Miller (1986) observes that 'it was not until the 1980s that the majority community [in the United Kingdom]...began to look at its own institutionalized racism... [with] clear implications for language policies' (291). Pilot searching for relevant studies indicated that research published before 1985 was not applicable to the topic of this review; therefore, only research published after 1980 was included.

### *3.2. Inclusion and exclusion criteria*

A set of criteria intended to determine the exclusion or inclusion of individual studies to the review was developed. The inclusion and exclusion criteria were applied at three levels throughout the systematic process in order to ensure that the papers maintained a high standard of relevancy: at the title level, at the abstract level, and at the article level.

Rationales for individual criteria are given in *Section 3.2.3*; any criterion with an associated justification or further explanation is marked by an asterisk (\*) in *Section 3.2.1* and *Section 3.2.2*.

### *3.2.1. Inclusion criteria*

Studies that adhered to the following criteria were *included* in the systematic map:

1. The research utilized an intervention or treatment design to investigate the acquisition or development of literacy or biliteracy in bilingual students.
2. The participants were of primary or elementary school age (3-11 years).
3. The study focused on participants who were bilingual and educated in a language other than their home language.\*
4. The research employed a quantitative design.
5. The intervention or treatment was implemented or administered through a school.
6. The research was available through OxLIP+ without purchase.
7. The research was published in English in the form of a journal article after 1980.

### *3.2.2. Exclusion criteria*

Studies that fit the following criteria were *excluded* from the systematic map:

1. The research followed an intervention or treatment design but did not have a control (non-treatment) group for comparison purposes.
2. The research investigated literacy acquisition or development in monolingual children, or did not distinguish between monolingual and bilingual children in the analysis of the intervention.
3. The research employed a qualitative design.
4. The journal name or article title indicated an interest in or focus on specific language impairments (SLIs), on students with behavioral, developmental, or linguistic disabilities, or on technology-based learning.\*
5. The control and treatment groups did not both include bilingual participants,\* or compared LM students against a non-LM population.

6. The researchers required the involvement of participants' caregivers as a necessary aspect of the intervention, including but not limited to parental consent forms\* or home-based supplemental activities.\*
7. A majority of the participants had behavioral, developmental, or linguistic disabilities that were not accounted for by their limited proficiency in the majority language, or a lack thereof.

### 3.2.3. Rationales for the inclusion and exclusion criteria

*Inclusion criterion 3:* Although sign language does not have a written form and, as such, deaf students in English-speaking environments learn to write in English as their first written language, studies examining the literacy of deaf students not educated in sign language—their native language—were considered to adhere to this criteria.

*Exclusion criterion 4:* While 'at-risk students' may refer to those students with a mental, physical, psychological, linguistic, social, or other impairment that would affect their achievement or performance in an academic setting, the designation may also denote students who are 'at risk' as a result of their language status, such as language minority children educated in a language that is not their home language. Therefore, studies referring to 'at-risk' students were not excluded as a result of exclusion criterion 4 until the nature of the mentioned 'risk' had been established as not related to the language status of the participants.

*Exclusion criterion 5:* While the participants in the treatment of intervention groups must necessarily be bilingual in order to adhere to inclusion criteria 3, the participants in the control group must also be bilingual in order to accurately act as a measure against which the researchers might gauge the progress of the treatment or intervention groups.

*Exclusion criterion 6a:* Any study that requires a parent or guardian to explicitly give consent for their child's participation necessarily excludes those children whose parents or guardians are uninterested in educational research or who are unable to provide consent due a language barrier, while parents or guardians who consent to such research may have external motivations for allowing their children to participate, skewing the generalizability of any results or conclusions drawn. Research conducted through a school, instructional policy, or other educational institution does not allow parents' or guardians' external

motivation for consenting or denying participation to obscure the results of such a study by removing it as a confounding variable. Research wherein the parental consent rate was 100% was not excluded on the understanding that the participants were not differentiated based on parental consent and thusly any conclusions were not distorted by the presence of parental consent as a confounding variable.

*Exclusion criterion 6b:* Ovando, Combs and Collier (2005) note that the extent to which parents are involved with the education of their children may affect academic achievement. Requiring a parent or guardian to contribute to their children's extracurricular learning, outside of a controlled intervention environment, entails relinquishing a measure of reliability given that the actual participation on the part of the non-student member is self-reported. Additionally, children with absentee, negligent, or busy guardians may not be given full advantage of the intervention program, producing skewed quantitative data.

### 3.3. *Electronic searching: Databases*

The database search consisted of entering a series of keywords, and their synonyms when applicable, in various combinations to produce a list of references relevant to the topic. The keywords and associated synonyms are listed in *Section 3.3.2*.

#### 3.3.1. *Databases*

The following databases were searched using the keywords found in *Section 3.3.2* and resulted in a total of 871 studies before the application of the inclusion and exclusion criteria and the removal of duplicate papers:

- British Education Index (BEI)
- Educational Resources Information Centre (ERIC – CSA)
- Linguistics and Language Behavior Abstracts (LLBA)
- PsychINFO
- Web of Science

The availability of the papers identified through the database search was subject to the subscriptions and permissions available through OxLIP+.

#### 3.3.2. *Keywords*

The following keywords, adjusting for spelling divergences when necessary, were employed to filter the relevant studies available through the database search:

- Program implementation / intervention / treatment
- Second language / foreign language / additional language
- Children at risk / minority group / language minority
- Bilingualism / biliteracy
- Literacy acquisition / development / skills
- Reading acquisition / development / skills / achievement
- Early childhood education

Due to the overlap among the studies located in each database, duplicate sets of some articles appeared in the search. Pilot searches with a selection of the above keywords ensured that they resulted in relevant studies when used in conjunction with each other. Table 3.1 (located at the end of this chapter) demonstrates the way in which the keywords were combined to locate relevant studies.

### *3.3.3. Application of keywords and inclusion and exclusion criteria*

122 unique papers remained from the database search after the duplicate results were eliminated prior to the abstract-level application of the inclusion and exclusion criteria. 51 of those studies adhered to the inclusion and exclusion criteria at the abstract level.

Table 3.2 (located at the end of this chapter) illustrates the application of the inclusion and exclusion criteria to the studies discovered through the database search. The numbers represent how many articles were identified through each stage of the searching process.

### *3.3.4. Systematic map*

A systematic map of the database search is located on at the end of this chapter.

## *3.4. Hand-searching: Journals*

After looking at the preliminary results of the database searching and examining the available journals archived via OxLIP+, 18 journals were singled out as the most likely to contain relevant studies, either by contributing to the results of the database search or by the themes implied by the title. Only those journals and archives available through OxLIP+ were examined so as to avoid incurring the cost of an individual subscription by the author of this review.

### *3.4.1. Journals*

The following journals were hand-searched for articles that appeared to adhere to the inclusion and exclusion criteria based on their titles (title-level application):

1. *Bilingualism* (1998-2011)
2. *The Bilingual Review/La revista bilingüe* (1999-2011)
3. *Developmental Psychology* (1980-2011)
4. *ELT Journal* (1980-2011)
5. *Foreign Language Annals* (1996-2011)
6. *Heritage Language Journal* (2003-2011)
7. *International Journal of Bilingual Education and Bilingualism* (1998-2011)
8. *International Journal of Bilingualism* (1999-2011)
9. *Journal of Educational Psychology* (1980-2011)
10. *Journal of Research in Reading* (1997-2011)
11. *Language Acquisition* (1990-2005)
12. *Language Learning* (1987; 1993-2011)
13. *Language and Literacy* (1999-2011)
14. *Modern Language Journal* (1997-2011)
15. *Reading in a Foreign Language* (1983-2011)
16. *Studies in Second Language Acquisition* (1997-2011)
17. *TESOL Quarterly* (1981-2007)
18. *The Canadian Modern Language Review/La revue canadienne des langues vivantes* (1980-2011)

The years of each journal's hand-searching were moderated by their electronic availability and by the seventh inclusion criteria (see *Section 3.1.1*) when issues archived prior to 1980 existed. After combining the studies identified during the hand-searching and database searching processes and eliminating any duplicate papers, no additional studies were found that fulfilled the inclusion and exclusion criteria, indicating that the parameters of the database searching were sufficient to identify relevant studies for this review.

#### 3.4.2. *Limitations of hand-searching*

While every effort was made to ensure that all relevant research was included in the hand-searching based on the criteria outlined in *Section 3.2*, the preliminary evaluation of each study was conducted based initially on the title of the article and its keywords, if available, and later on the abstract of the article. Titles and abstracts may not have contained all relevant information and so any validation of the systematic process must necessarily

account for a degree of human error, specifically that stemming from any lack of transparency in the literature.

### *3.5. Authenticity of systematic process*

In order to ensure that the systematic process had indeed uncovered all of the appropriate studies, the works of several researchers known for their interest in the field of bilingual education were examined. Any relevant material uncovered during the authentication process that was not identified by the original keyword searching resulted in the addition, substitution, or modification of keywords to the search, until the systematic map recognized all pertinent research as such during the search process.

For example, the term ‘early childhood education’ was substituted for ‘elementary / primary education’ as an alternative label for studies whose population consisted of elementary- or pre-elementary-aged children in order to restrict the search to comply with inclusion criterion (see *Section 3.2.1.*) when the latter resulted in mixed-aged populations and did not result in studies focusing on children in nursery school or pre-kindergarten.

### *3.6. Verification of results: Abstract level*

In order to verify the suitability of the 122 studies that were produced by the initial database-searching, the abstracts of each study were read and analyzed according to the inclusion and exclusion criteria (abstract-level). 51 studies remained that, at the abstract level, concerned intervention studies focusing on the acquisition of literacy in language minority children educated in a second language and were considered relevant to the focus of this review.

### *3.7. Confirmation of inclusion and exclusion criteria: Article level*

The 51 studies that were identified by the database and journal searching methods as adhering to the inclusion and exclusion criteria according to their titles and abstracts were read in their entirety to ensure complete adherence to each criterion during the verification process (see *Section 3.5*). 44 non-relevant studies were discarded, leaving 7 for the in-depth analysis.

### *3.8. In-depth review*

The studies identified in *Section 3.6* as appropriate for the topic were subject to the following additional in-depth inclusion criteria in order to ensure that only the studies most relevant to the review were included:

1. The study consisted of an intervention or treatment program that focused on literacy and literacy-related skills.
2. The intervention or treatment program did not involve a focus on any non-literacy-related skill.
3. The intervention or treatment program did not involve the participants' caregivers or extracurricular input thereof.

All 7 studies adhered to the in-depth inclusion criteria and remained to be considered for inclusion in the in-depth analysis subject to the application of the Weight of Evidence criteria in *Section 3.9*.

### *3.9. Weight of Evidence*

The studies that passed the inclusion and exclusion criteria in *Section 3.2.1*, *Section 3.2.2*, and *Section 3.8* at the article level were subjected to a Weight of Evidence (WoE) evaluation in order to assess their external and internal validity. The WoE criteria were adapted from the EPPI-Centre guidelines for WoE analysis (see Gough, 2007; 'Quality and Relevance Appraisal') and were:

1. *Validity of hypotheses or research questions*: How appropriate were the stated research questions or hypotheses of the study with regards to its analysis and conclusions?
2. *Suitability of design and measurements*: How appropriately was the research conducted?
3. *Relevance to the topic investigated by this review*: How germane to the research questions of this review was the study?
4. *Efficacy*: How efficient was the intervention in terms of its outcome?

Only those 5 papers that were rated at least moderately relevant for each of the WoE criteria were included in the in-depth analysis. *Appendix A* shows the ratings assessment of the 7 articles that adhered to the inclusion and exclusion criteria after the in-depth criteria analysis outlined in *Section 3.8*.

Of the 5 papers that adhered to all of the inclusion and exclusion criteria and were rated by the author as moderately relevant for each WoE criterion, 2 (40%) were independently assessed by a second source in order to determine the validity of the WoE requirements. A substantial level of inter-rater reliability was established using statistical analysis (Kappa = 0.750; see Landis & Koch, 1977) to validate the WoE criteria and evaluation process (see *Appendix B*).

**TABLE 3.1: KEYWORD APPLICATION STRATEGY**

OR		OR		OR
<ul style="list-style-type: none"> <li>• Intervention</li> <li>• Treatment</li> <li>• Program* implementation</li> </ul>	AND	<ul style="list-style-type: none"> <li>• Second language</li> <li>• Foreign language</li> <li>• Additional language</li> <li>• Children at risk</li> <li>• Minority group</li> <li>• Language minority</li> <li>• Early childhood education</li> <li>• English language learners</li> </ul>	AND	<ul style="list-style-type: none"> <li>• Bilingualism</li> <li>• Biliteracy</li> <li>• Literacy development/ acquisition/ skills</li> <li>• Reading development/ acquisition/ skills</li> <li>• Reading achievement</li> </ul>

**TABLE 3.2: DATABASE SEARCHING: APPLICATION OF INCLUSION AND EXCLUSION CRITERIA**

		BEI	ERIC	LLBA	Web of Science	PsycINFO	Total
Keyword search		70	228	114	341	118	871
Application of Inclusion and Exclusion Criteria	Title	38	69	26	41	42	232
	Unique records	36	59	12	25	42	175
	<i>Available:</i>	23	43	5	16	35	122
	Abstract						51
	Article						9
<i>TOTAL</i> after application of in-depth criteria							7
<i>TOTAL</i> after WoE criteria ratings: inclusion in in-depth analysis							5

**FIGURE 1: SYSTEMATIC MAP**

