56. The Crown reserves its rights to seek compensation from the Owner and reserves all its remedies at common law and equity if the Owner breaches any warranty granted in clauses 50 to 55 (inclusive) of this Schedule D.

General

- 57. The parties must each sign, execute, procure, pass and do all such further documents, acts, matters, resolutions and things as may be necessary or desirable for effecting the transactions contemplated by this Agreement.
- 58. The agreements, obligations and warranties contained in the Agreement will not merge on completion of the transactions contemplated by them but will remain in full force until satisfied.
- 59. The Owner hereby acknowledges and agrees that until accepted in writing, this document is only an offer to sell to the Crown and its acceptance will not be anticipated and the Owner will not enter into any consequential commitments in reliance of this offer being accepted by the Crown.
- 60. Any facsimile or copy of this Agreement transmitted by e mail (including any copy transmitted by e mail or facsimile copy of any document evidencing the execution of this Agreement by either party) may be relied upon by the other party as though it was an original copy.
- 61. The parties acknowledge and agree that where any obligation or agreement in the Agreement remains unperformed at the Settlement Date then that obligation or agreement, notwithstanding any rule of law or equity to the contrary, will enure until fully discharged by performance and in no circumstances whatsoever will merge upon settlement.
- 62. The terms and conditions set out in this Agreement and any approvals and consents in writing provided for in this Agreement and given prior to execution, contain the entire agreement as concluded between the parties. The Owner acknowledges that the Crown makes no representation or warranty other than as are expressly set out and described in this Agreement. The Owner warrants and undertakes to the Crown that it has entered into this Agreement entirely in reliance on its own judgement and inquiries and confirms that, in entering into this Agreement, the Owner has not relied on any statement, representation, warranty, condition, promise or undertaking made by or on behalf of the Crown in the course of communications or negotiations, whether express or implied by conduct, prior to or during the making of this Agreement and not expressly set out in this Agreement.

Obligations on Disposal of Owner's Land

(a)

- 63. The Owner must not dispose of (whether by sale, gift, assignment, transfer or change of shareholding);
 - its effective control of that part of the balance of the Owner's Land over which the Owner Works are to be constructed prior to Practical Completion of the Owner Works and discharge in full of all the Owner's obligations under this Agreement; and/or
 - (b) any part of the balance of the Owner's Land not included in clause 63(a) above prior to the Ministry obtaining the designation over the Required Land, and operating the Required Land as a school site;

without first obtaining the Crown's prior written consent, which will not be unreasonably withheld if:

(a) where clause 63(a) applies:

- the Crown is satisfied that the assignee or transferee is a respectable, solvent and reputable party with sufficient standing and experience to fulfill the obligations of the Owner under this Agreement; and,
- (ii) the assignee or transferee enters into a deed of covenant in favour of the Crown whereby the assignee or transferee acknowledges the rights and interests of the Crown pursuant to this Agreement and agrees to pay, perform and observe the obligations of the Owner under this Agreement. The entry into such deed of covenant by such transferee or assignee will not release the Owner from its obligations under this Agreement.
- (c) Where clause 63(b) applies, the assignee or transferee enters into a deed of covenant in favour of the Crown whereby the assignee or transferee acknowledges the rights and interests of the Crown pursuant to clause 45 of this Schedule D, and agrees to pay, perform and observe the obligations contained in clause 46 of this Schedule D. The entry into such deed of covenant by such transferee or assignee will not release the Owner from its obligations under this Agreement.
- 64. Clause 63 does not apply to any part of the Owner's Land vested as reserve, road, or in which a utility company has an interest for the purposes of locating utilities, or to any part of the balance of the Owner's Land where, at the date of this Agreement, the Owner has already entered into agreements for Sale and Purchase with purchasers being the land detailed in Schedule I.

Dispute Resolution - Matters Relating to the Owner Works and Retention Sum

(c)

- 65. If an issue arises between the parties in any way relating to or concerning the Owner Works (but excluding any matter arising under clause 35(a) of this Schedule D) or the release of the Retention Sum and the parties are unable to resolve the issue within 10 Working Days of the issue arising, the issue shall be referred for determination to an expert who is relevantly qualified and experienced, and recognised as an expert, in relation to the subject matter of the issue. The parties shall use all reasonable endeavours to assist the expert in resolving the issue.
- 66. The expert shall be appointed by agreement of the parties or, in the absence of agreement within 10 Working Days, upon the request of either party, by:
 - the President of the Institution of Professional Engineers New Zealand (or any successor organisation), in respect of any issue relating to engineering or construction matters; or
 - (b) the President of the New Zealand Institute of Surveyors (or any successor organisation) in respect of any issue relating to survey matters; or
 - the President of the New Zealand Institute of Valuers (or any successor organisation), in respect of any issue relating to valuation matters; or
 - (d) the President of the New Zealand Law Society (or any successor organisation), in respect of any other issue.
- 67. The expert shall act as an expert and not as an arbitrator, and the expert's opinion shall be final and binding on the parties, except in the case of manifest error or fraud.
- 68. The parties shall each bear their own costs in respect of the issue referred to the expert for determination and shall each pay half of all costs incurred by the expert, unless the expert (at his or her discretion) determines otherwise. Any such determination as to costs shall be binding on the parties.

Dispute Resolution - General and Interpretation Issues

- 69. If any dispute arises between the parties concerning the interpretation of this Agreement or relating to any matter arising under the Agreement, apart from those matters in respect of which a specific dispute resolution process has been provided for under this Agreement, the parties will actively and in good faith enter into negotiations with each other with a view to a speedy resolution of such dispute or disputes.
- 70. If the parties cannot resolve a dispute within fifteen (15) Working Days of any dispute arising then, unless otherwise expressly provided in this Agreement, they will without prejudice to any other right, explore whether such dispute can be resolved by agreement between them using informal dispute resolution techniques such as mediation. The rules governing any such technique if adopted will be agreed between the parties or as selected by the organisation known as "The Resolution Institute".
- 71. If the parties cannot agree on any dispute resolution technique within a further fifteen (15) Working Days of any dispute being considered for referral by both parties to any informal dispute resolution technique under clause 69 above then the dispute shall be settled by reference to arbitration. Except as otherwise expressly provided in this Agreement the reference shall be to a single arbitrator if one can be agreed upon, or to two arbitrators (one to be appointed by each party) and their umpire (appointed by them prior to their arbitration), such arbitration to be carried out in accordance with the Arbitration Act 1996 and the substantive law of New Zealand.
- 72. The parties will co-operate to ensure the expeditious conduct of any arbitration. In particular, each party will comply with any reasonable time limits sought by the other for settling terms of reference, interlocutory matters and generally all steps preliminary and incidental to the hearing and determination of the proceedings.

Notices

- 73. All notices and other communications required or permitted under this Agreement shall be in writing and shall be delivered personally or by electronic mail (email) to the address set out below (or such address as may be directed in writing from time to time by the parties).
- 74. Where notice or communications are sent by email notice shall be deemed served when sent to a notified email address and receipt acknowledged orally by the recipient or by writing including return email.
- 75. Where notice or communications are personally delivered notice shall be deemed served when received at the recipients notified address prior to 4pm on any working day. Where delivered after 4pm notice shall be deemed served on the next working day.

To the Owner:

Hughes Development Limited (Attention: Jake Hughes) 21 Worcester Boulevard PO Box 848, Christchurch 8014 Email: jake@hughesdevelopments.co.nz

To the Owner's Solicitor:

Parry Field Lawyers PO Box 8020, Riccarton, Christchurch, 8440 Email: timrankin@parryfield.com

To the Ministry: Ministry of Education (Attention: Deb Taylor) 48 Hereford Street Christchurch Email: <u>deb.taylor@education.govt.nz</u>

To the Crown:

Manager Clearances Crown Property Management Land Information New Zealand PO Box 5501 Wellington 6145 Email: <u>Clearances@linz.govt.nz</u>

To the Crown's accredited supplier:

The Property Group (Attention: Sarah Oakley) Level 2, Lion House 169 Madras Street Christchurch 8011 Email: <u>soakley@propertygroup.co.nz</u>

To the Crown's solicitor:

C/- EIS Legal, Ministry of Education (Attention: Tracy Finlayson) Mātauranga House Level 1 33 Bowen Street Wellington 6011 Email: <u>Tracy.Finlayson@education.govt.nz</u>

- 76. All parties may amend their contact details above by notice to all the other parties.
- 77. A copy of all notices or communications which are required under this Agreement to be given to the Crown must also be given to the Ministry and the Crown's solicitors.

Electronic Execution

- 78. The parties mutually acknowledge and agree that:
 - (a) this Agreement may be executed in two or more counterparts, all of which will together be deemed to constitute one and the same agreement. A party may enter into this Agreement by signing a counterpart copy and sending it to the other party, including by facsimile or email.
 - (b) the production of an emailed copy or copies of the Agreement signed by all parties shall be deemed to be sufficient to satisfy the requirements of the Property Law Act 2007.
 - (c) in the case of email, any notice or document is deemed to have been received when acknowledged by the party by return email or otherwise in writing, except that return emails generated automatically shall not constitute an acknowledgement.
 - (d) in accordance with the Contract and Commercial Law Act 2017, the parties agree that any notice or document that must be given in writing by one party to the other may be given in electronic form and by means of electronic communication.

23

Confidentiality

eleased under the

79. The parties acknowledge that the terms of this Agreement are confidential. Subject to the Crown's obligations under the Official Information Act 1982, the parties agree that they will not at any time, unless authorised in writing by the other party, directly or indirectly, disclose to any third party (except to a party's officers, employees or professional advisers and only on the basis that the third party is aware of the confidential nature of this Agreement and will not disclose such information) any information in respect of this Agreement including the terms and Compensation payable.

sticial

24

5

Released Under the Official Information AC 25





Title Plan - DP 557037

Survey Number	DP 557037			$\mathbf{\Lambda}$			
Surveyor Reference	H20242 Selwyn Rd						
Surveyor	Nicholas Johan Jagvik						
Survey Firm	Davie Lovell-Smith Ltd	e constante de la facta de la casa de la cas	Charles Co.				
Surveyor Declaration	I Nicholas Johan Jagvik, (a) this dataset provided Cadastral Survey Act 200 (b)the survey was under Declared on 23 Mar 202	being a licensed cadastral surveyor, co by me and its related survey are accur 22 and the Rules for Cadastral Survey taken by me or under my personal dire 1 09:11 AM	ertify that: rate, correct and in 2010, and ection.	accordance with th			
Survey Details			<u>x</u>				
Dataset Description	Lots 350, 355, and 1000- 341771, Lots 1 & 3 DP 4	1003 being subdivision of Lot 1 DP 6 41634 and Lots 1 & 2 DP 479375	0892, Lot 2 DP 6	3632, Lots 1 & 2 DI			
Status	Deposited		•				
Land District	Canterbury	Survey Class	Class A				
Submitted Date	23/03/2021						
		Deposit Date	31/03/2021				
Territorial Authoritie	s	0					
Selwyn District							
Comprised In		No.					
RT CB36A/800	X						
RT 171911							
RT 171912							
RT 549973	XO						
RT 667881							
RT 667882							
Created Parcels	9						
Parcels		Parcel Intent	Area	RT Reference			
Lot 1003 Deposited Pl	lan 557037	Fee Simple Title	0.0833 Ha	977883			
Lot 1000 Deposited P	lan 557037	Fee Simple Title	10.1837 Ha	977880			
Lot 350 Deposited Pla	m 557037	Fee Simple Title	0.0005 Ha	985683			
Lot 355 Deposited Pla	n 557037	Fee Simple Title	0.0002 Ha	985684			
Lot 1001 Deposited P	lan 557037	Fee Simple Title	11.7932 Ha	977881			
Lot 1002 Deposited P	lan 557037	Fee Simple Title	13.1970 Ha	977882			
Total Area			35.2579 Ha				

Page 1 of 5



Davie, Lovell-Smith Ltd 116 Wrights Road, Addington PO Box 679, Christchurch 8140 New Zealand Telephone (03) 379-0793, E-mail: office@dls.co.nz

Surveyor Reference: H20242 Selwyn Rd

eleased

DP 557037

SCHEDULE OF EASEMENTS IN GROSS

Purpose	Servient Tenement (Burdened Land)	Shown	Grantee
Right to convey electricity in gross	Lot 1003	Lot 1003	Orion New Zealand Limited
Right to convey telecommunications in gross	Lot 1003	Lot 1003	Enable Networks Limited

PART IVA CONSERVATION ACT 1987

Lots 355 & 1003 and part of Lots 1000 & 1001, formerly Lot 1 DP 341771 (RT 171911), Lot 2 DP 341771 (RT 171912) and Lot 1 DP 60892 (RT CB36A/800), are subject to Part IV A Conservation Act 1987

1







Schedule F: Specifications of Owner Works

	Construction of roading and installation of services in accordance with Roading			
-	Plans in Schedule G of this Agreement.			
2.	Any works identified pursuant to clause 47(b), 47(c) and 48(b) of Schedule D	Within 6 months of being identified and agreed pursuant to clause 48(b) of Schedule D	\$300,00	
Req	uirements for Owners Works	XC)	
1.	Roads and footpaths to be constructed to minimum of Council Standard			
2.	Fibre, electricity, potable water supply, including water supply for fire-fighting to mee PAS 4509:2008 FW 2 classification requirements (total firefighting water flow of required while maintaining a minimum residual pressure in the water supply line of kPa), waste water (including sanitary sewer) discharge services to be installed connection points in the Required Land at locations negotiated with the Ministr operational, connected and commissioned.			
3.	Roading to be provided to the site bounda G.	aries as illustrated on the conce	pt plan at S	
-				

Ø 26

eleased Under the Official Information A






















































Schedule H: Engineer's Deed of Covenant

Deed of covenant dated

Parties

(Owner)

(Engineer)

HER MAJESTY THE QUEEN ACTING BY AND THROUGH THE SECRETARY FOR EDUCATION (Crown)

Introduction

- A. The Owner and the Crown are parties to an agreement entered into between the Owner and the Crown dated [#] in respect of the sale of the Required Land (as defined in the Agreement) to the Crown (Agreement). Pursuant to the Agreement the Owner will construct and deliver certain Owner Works (as defined in the Agreement) in accordance with the Agreement.
- B. The Engineer has been supplied with a copy of the Agreement and has been engaged by the Owner to carry out the role of Owner's Engineer (as defined in the Agreement).
- C. The role of Owner's Engineer under the Agreement is, independently of either the Owner or the Crown, to fairly and impartially make decisions entrusted to the Engineer under the Agreement including (without limitation) as to:
 - 1. the carrying out and completion of the Owner Works;
 - 2. determining disputes arising under the Agreement,

(the Role).

- D. The Owner has agreed to procure this deed of covenant from the Engineer pursuant to the Agreement.
- E. The Engineer has agreed, in consideration of its engagement, to provide the covenants set out in this Deed in favour of the Crown.

Covenants

The Engineer warrants to the Crown that it will perform the Role independently of the Owner, and fairly and impartially and with all due care and professional skill, including when:

- (a) executing certificates required under the Agreement; and
- (b) considering information and dealing with parties in contemplation of, during or following a dispute between the Owner and the Crown in relation to the Role.

2021

- 2. The Engineer shall exercise the Role without delay.
- 3. References in this Deed to the Crown includes the Ministry of Education and the Crown may exercise its rights and comply with its obligations under this Deed by and through the Ministry of Education.

Execution

Signed for and on behalf of **[INSERT]** as Owner in the presence of:

Signature of witness

Name of witness

Occupation of witness

City/town of residence

Signed by [*insert name*] as Engineer in the presence of:

Signature of witness

Name of witness

Occupation of witness

City/town of residence

Signed by and on behalf of HER MAJESTY THE QUEEN by [insert Name] acting pursuant to a delegation from the Secretary for Education in the presence of:

Signature of witness

Name of witness

Occupation of witness

City/town of residence

Signature of authorised person

Name of authorised person

Signature of authorised person

Name of authorised person

Signature

[Insert name]

Signature of authorised person

Name of authorised person

Schedule I: Land sold pursuant to Existing Agreements for Sale and Purchase and excluded from the Owner's covenant under clause 63 (clause 64)

- Lots 37 to 57 (inclusive)
- Lots 59 to 82 (inclusive)
- Lots 83 to 106 (inclusive)
- Lots 108 to 117 (inclusive)
- Lots 243 to 258 (inclusive)
- Lots 261 to 266 (inclusive)
- Lots 275 and 276
- Lots 277 to 286 (inclusive)

eleasedunderth

As identified on plan titled 'South East Faringdon Stages 4, 5 & 7 Overall Plan' drawing no: H.20242.STG4,5,7 E01.0 R6' contained in Schedule G.

30

5

VISIBLE LEARNING A SYNTHESIS OF OVER 800 META-ANALYSES RELATING TO ACHIEVEMENT



Visible Learning

elede

This unique and ground-breaking book is the result of 15 years' research and synthesises over 800 meta-analyses relating to the influences on achievement in school-aged students. It builds a story about the power of teachers and of feedback, and constructs a model of learning and understanding.

Visible Learning presents research involving many millions of students and represents the largest ever collection of evidence-based research into what actually works in schools to improve learning. Areas covered include the influences of the student, home, school, curricula, teacher, and teaching strategies. A model of teaching and learning is developed based on the notion of visible teaching and visible learning.

A major message within the book is that what works best for students is similar to what works best for teachers. This includes an attention to setting challenging learning intentions, being clear about what success means, and an attention to learning strategies for developing conceptual understanding about what teachers and students know and understand.

Although the current evidence-based fad has turned into a debate about test scores, this book is about using evidence to build and defend a model of teaching and learning. A major contribution to the field, it is a fascinating benchmark for comparing many innovations in teaching and schools.

John Hattie is Professor of Education and Director of the Visible Learning Labs, University of Auckland, New Zealand.

eleased under the Orticial Information A

Visible Learning Stion A synthesis of over 800 meta-analyses relating to achievement eleased under the official



ration A

First published 2009 by Routledge 2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada by Routledge 270 Madison Avenue, New York, NY 10016

Routledge is an imprint of the Taylor & Francis Group, an informa business

This edition published in the Taylor & Francis e-Library, 2008.

"To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk."

© 2009 John A. C. Hattie

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

British Library Cataloguing in Publication Data A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data Hattie, John. Visible learning: a synthesis of meta-analyses relating to achievement/John A. C. Hattie. p. cm.

Includes bibliographical references.

I. Learning Longitudinal studies. 2. Teaching—Longitudinal studies. 3. Effective teaching Longitudinal studies. 4. Teacher effectiveness—Longitudinal studies. I. Title.

LB1060.H388 2008 370.15'23—dc22

2008021702

ISBN 0-203-88733-6 Master e-book ISBN

ISBN10: 0-415-47617-8 (hbk) ISBN10: 0-415-47618-6 (pbk) ISBN10: 0-203-88733-6 (ebk)

ISBN 13: 978-0-415-47617-1 (hbk) ISBN 13: 978-0-415-47618-8 (pbk) ISBN 13: 978-0-203-88733-2 (ebk)

Contents

	i chi	
	List of tables	V1
	List of figures	V11
	A chnowladoments	V111
		А
1	The challenge	1
2	The nature of the evidence: a synthesis of meta-analyses	7
3	The argument: visible teaching and visible learning	22
4	The contributions from the student	39
5	The contributions from the home	61
6	The contributions from the school	72
7	The contributions from the teacher	108
8	The contributions from the curricula	129
9	The contributions from teaching approaches—part I	161
10	The contributions from teaching approaches—part II	200
11	Bringing it all together	237
	Appendix A: the meta-analyses by topic Appendix B: the meta-analyses by rank order Bibliography Index	263 297 301 375

Tables

		R
	• (
2.1	Average effect for each of the major contributors to learning	18
4.1	Summary information from the meta-analyses on the contributions of from the student	39
5.1	Summary information from the meta-analyses on the contributions	61
6.1	Summary information from the meta-analyses on the contributions	74
6.2	Synthesis of meta-analyses and major studies reducing class size from	/4
7.1	25 to 15 Summary information from the meta-analyses on the contributions	87
8.1	from the teacher Summary information from the meta-analyses on the contributions	109
9.1	from the curricula	130
2.1	from teaching approaches	162
9.2	Relation between goal difficulty and performance	165
9.3	Difficult compared to "do your best" goals	165
9.4	Relation of self-efficacy to goal attainment	166
9.5	Various metacognitive strategies and the effect sizes (Lavery, 2008)	190
10.1	Summary information from the meta-analyses on the contributions	
	from teaching approaches	201
10.2	Effect sizes for various teaching strategies (from Seidel & Shavelson, 2007)	202
10.3	Effect sizes for various teaching strategies (from Marzano 1998)	203
10.4	Summary of effects from comprehensive teaching reforms	
	(Borman <i>et al.</i> , 2003)	216
10.5	Summary of effects from computer-based instruction	222
10.6	Summary of effects from computers as substitute and as supplement	222
10.7	to the teacher	223
10.7	Summary of effects from using computers with the same or a	222
10.0	different teacher	223
10.8	Summary of major uses of computers in classrooms	224
11.1	Effect sizes for teacher as activator and teacher as facilitator	243
11.2	Effect sizes from teaching or working conditions	244

Figures

		V
		<u>, '</u>
		Θ^{-}
1.1	Percentage of responses as to the claimed influences on student	,
2.1	An achievement continuum	5
2.1	Distribution of effect sizes across all meta-analyses	16
2.3	Number of meta-analyses above and below the h-point	19
2.4	A typical barometer of influence	19
2.5	Funnel plot of the effect size and sample size from each meta-analysis	21
7.6	Effect sizes for nine teacher-student relationship variables	119
8.10	The percentage of students performing below, at, and above	
	expectation in each year level	141
8.24	Higgins <i>et al.</i> four-part thinking model	156
9.9	A model of feedback	176
10.18	The number of computer-based meta-analyses and their	
10.10	overall effect size	220
10.19	Relation between effect sizes for computer-based instruction	221
11 1	A model of Visible terrbing. Visible learning	221
11.1	The means for the National Board cartified teachers (NIBCTs) and	230
11.2	non-National Board certified teachers (non-NBCTs) and the	
	effect size of the difference between these two groups	260
11.3	Percentage of student work from NBCTs and non-NBCTs	200
	classified as surface or deep learning	260
	S	
	<u>0</u>	
V		
2		

Preface

Elliott is my hero. On his fifth birthday he was diagnosed with leukemia, and this past year has been his *annus horribilis*. On the day of the diagnosis, it was impressive to see the medical team immediately begin interventions. While they aimed to make Elliott stable, the diagnosis regime burst into action. They knew which tests were needed to make the correct diagnosis and when they were satisfied with the initial diagnosis they immediately moved to interventions. Thus began a year of constant monitoring and feedback to the medical team about Elliott's progress. All throughout they collected evidence of progress, they knew what success looked like, and kept all informed about this evidence. Elliott went through many ups and downs, lost his hair (as did I when he gave me a No. 1 cut as his Christmas present, although I drew a line when he asked to shave my eyebrows off as well), and had daily injections in the front of his legs, but he never balked, and throughout the treatment maintained his sparkly personality. The family was never in the dark about what was happening, books were provided, sessions offered, and support for treatment was excellent. The messages in this book owe a lot to Elliott.

This book started in Gil Sax's office in 1990 searching and coding meta-analyses. Motivation to continue the search was inspired by Herb Walberg, and continued in Perth in Australia, North Carolina in the US, and finished here in Auckland in New Zealand. It is a journey that has taken 15 years. The messages have been questioned, labelled provocative, liked, and dismissed, among other more positive reactions. The typical comments are: "the results do not mirror my experience", "why have you not highlighted my pet method", "you are talking about averages and I'm not average", and "you are missing the nuances of what happens in classrooms". There are many criticisms and misunderstandings about what I am *and* am not saying.

So let me start with what this book is not.

It is not a book about classroom life, and does not speak to the nuances and details of what happens within classrooms. Instead it synthesizes research based on what happens in classrooms; as it is more concerned with main effects than interactions. Although I have spent many hundreds of hours in classrooms in many countries, have observed, interviewed, and aimed to dig quite deeply into the nuances of classrooms, this book will not show these details of class living.

2 It is *not* a book about what cannot be influenced in schools—thus critical discussions about class, poverty, resources in families, health in families, and nutrition are not included—but this is NOT because they are unimportant, indeed they may be more important than many of the influences discussed in this book. It is just that I have not included these topics in my orbit.

- 3 It is *not* a book that includes qualitative studies. It only includes studies that have used basic statistics (means, variances, sample sizes). Again, this should not mean qualitative studies are not important or powerful but just that I have had to draw some lines around what can be accomplished over a 15-year writing span.
- 4 It is *not* a book about criticism of research, and I have deliberately not included much about moderators of research findings based on research attributes (quality of study, nature of design) again not because these are unimportant (my expertise is measurement and research design), but because they have been dealt with elsewhere by others (e.g., Lipsey & Wilson, 1993; Sipe & Curlette, 1996a, 1996b).

Rather this is a book about synthesizing many meta-analyses. It is based on over 50,000 studies, and many millions of students—and this is a cut down version of what I could have included as I also collected studies on affective and physical outcomes and on many other outcomes of schooling. I occasionally receive emails expressing disbelief that I have had the time to read so many studies. No, I have not read all primary studies, but as will be seen I have read all meta-analyses, and in some cases many of the primary studies. I am an avid reader, thoroughly enjoy learning the arts of synthesizing and detecting main ideas, and want to create explanations from the myriad of ideas in our discipline. The aim of this book is not to overwhelm with data—indeed my first attempt was discarded after 500 pages of trenchant details; who would care about such details? Instead this book aims to have a message, a story, and a set of supporting accounts of this story.

The message about schools is a positive one. So often when talking about the findings in this book, teachers think I am attacking them as below average, non-thinking, boring drones. In New Zealand, for example, it is clear to me why we rank in the top half-dozen nations in reading, mathematics, and science—we have a nation of excellent teachers. They exist and there are many of them. This book is a story of many real teachers I have met, seen, and some who have taught my own boys. Many teachers already think in the ways I argue in this book; many are seeking to always improve and constantly monitor their performances to make a difference to what they do; and many inspire the love of learning that is one of the major outcomes of any school. This is not a book claiming that teachers are below par, that the profession is terrible, and that we all need to "put in more effort and do better". Nearly all studies in the book are based on real students in front of real teachers in real schools and that so many of the effects are powerful is a testament that excellence is happening. The major message is that we need a barometer of what works best, and such a barometer can also establish guidelines as to what is excellent-too often we shy from using this word thinking that excellence is unattainable in schools. Excellence is attainable: there are many instances of excellence, some of it fleeting, some of it aplenty. We need better evaluation to acknowledge and esteem it when it occurs—as it does.

Acknowledgments

eleas

There are so many who have contributed to the data, the book, and the message, and who have provided feedback over these past 15 years: Nola Purdie, Krystoff Krawowski, Richard Fletcher, Thakur Karkee, Earl Irving, Trisha Lundberg, Lorrae Ward, Michael Scriven, Richard Jaeger, Geoff Petty, and Russell Bishop. I am especially indebted to Janet Rivers for her attention to the details and to Debbie Waayer for her remarkable skills in finding articles, referencing, and data skills, and in ensuring that I completed this book. Others have been critics and this is among the more welcome contributions for any author: Lexie Grudnoff, Gavin Brown, Adrienne Alton-Lee, Christine Rubie-Davis, Misty Sato, David Moseley, Heidi Leeson, Brian Marsh, Sandra Frid, Sam Stace, and John Locke. I particular thank Gene Glass for his development of meta-analysis that allowed me and many others to stand on his shoulders to peer into what makes a difference to teaching and learning.

But most of all I thank my family—they have endured this book, shaped the many versions of the message, and provided the feedback that only a loving family can give. Unlike most children who are asked about their day at school each night at the dinner table, my boys have endured the same interrogation every night of their school years: What feedback did you receive about your learning today? Thanks to my boys—Joel, Kyle, Kieran, Billy, Bobby, and Jamie—you are my inspirations for living. And most of all to Janet—the one who has given unconditional positive regard through the ups and downs of moving a family across many countries, putting up with "yet another study", and being the love of my life. The size of your effect on my life exceeds any reported in this book.

The challenge

In the field of education, one of the most enduring messages is that "everything seems to work". It is hard to find teachers who say they are "below average" teachers, and everyone (parent, politician, school leader) has a reason why their particular view about teaching or school innovation is likely to be successful. Indeed, rhetoric and game-play about teaching and learning seems to justify "everything goes". We acknowledge that teachers teach differently from each other; we respect this difference and even enshrine it in terms like "teaching style" and "professional independence". This often translates as "I'll leave you alone, if you leave me alone to teach my way."While teachers talk to their colleagues about curriculum, assessment, children, and lack of time and resources, they rarely talk about their teaching, preferring to believe that they may teach differently (which is acceptable provided they do not question one another's right to teach in their particular ways). We pass laws that are more about structural concerns than about teaching concerns: such as class size, school choice, and social promotion, as if these are clear winners among the top-ranking influences on student learning. We make school-based decisions about ability grouping, detracking or streaming, and social promotion, again appealing to claims about influences on achievement. For most teachers, however, teaching is a private matter; it occurs behind a closed classroom door, and it is rarely questioned or challenged. We seem to believe that every teacher's stories about success are sufficient justification for leaving them alone. We will see throughout this book that there is a good reason for acknowledging that most teachers can demonstrate such success. Short of unethical behaviors, and gross incompetence, there is much support for the "everything goes" approach. However herein lies a major problem.

It is the case that we reinvent schooling every year. Despite any successes we may have had with this year's cohort of students, teachers have to start again next year with a brand new cohort. The greatest change that most students experience is the level of competence of the teacher, as the school and their peers typically are "similar" to what they would have experienced the previous year. It is surely easy to see how it is tempting for teachers to re-do the successes of the previous year, to judge students in terms of last year's cohort, and to insist on an orderly progression through that which has worked before. It is required of teachers, however, that they re-invent their passion in their teaching; they must identify and accommodate the differences brought with each new cohort of students, react to the learning as it occurs (every moment of learning is different), and treat the current cohort of students as if it is the first time that the teacher has taught a class—as it is for the students with this teacher and this curricula.

As will be argued throughout this book, the act of teaching reaches its epitome of

success after the lesson has been structured, after the content has been delivered, and after the classroom has been organized. The art of teaching, and its major successes, relate to "what happens next"—the manner in which the teacher reacts to how the student interprets, accommodates, rejects, and/or reinvents the content and skills, how the student relates and applies the content to other tasks, and how the student reacts in light of success and failure apropos the content and methods that the teacher has taught. Learning is spontaneous, individualistic, and often earned through effort. It is a timeworn, slow and gradual, fits-and-starts kind of process, which can have a flow of its own, but requires passion, patience, and attention to detail (from the teacher and student).

So much evidence

The research literature is rich in recommendations as to what teachers and schools should do. Carpenter (2000), for example, counted 361 "good ideas" published in the previous ten years of Phi Delta Kappan (e.g., Hunter method, assertive discipline, Goals 2000, TQM, portfolio assessment, essential schools, block scheduling, detracking, character education). He concluded that these good ideas have produced very limited gains, if any. Similarly, Kozol (2005, p. 193) noted that there have been "galaxies of faded names and optimistic claims," such as "Focus Schools", "Accelerated Schools", "Blue Ribbon Schools", "Exemplary Schools", "Pilot Schools", "Model Schools", "Quality Schools", "Magnet Schools", and "Cluster Schools"-all claiming they are better and different, with little evidence of either. The research evidence relating to "what works" is burgeoning, even groaning, under a weight of such "try me" ideas. Most are justified by great stories about lighthouse schools, inspiring principals and inspiring change agents, and tales of wonderful work produced by happy children with contented parents and doting teachers. According to noted changetheory expert, Michael Fullan, one of the most critical problems our schools face is "not resistance to innovation, but the fragmentation, overload, and incoherence resulting from the uncritical and uncoordinated acceptance of too many different innovations (Fullan & Stiegelbauer, 1991, p. 197). Richard Elmore (1996) has long argued that education suffers not so much from an inadequate supply of good programs as from a lack of demand for good programs—and instead we so often supply yet another program rather than nurture demand for good programs.

There is so much known about what makes a difference in the classroom. A glance at the journals on the shelves of most libraries, and on web pages, would indicate that the state of knowledge in the discipline of education is healthy. The worldwide picture certainly is one of plenty; we could have a library solely consisting of handbooks about teaching, most of which cannot be held in the hand. Most countries have been through many waves of reform, including new curricula, new methods of accountability, reviews of teacher education, professional development programs, charter schools, vouchers, and management models. We have blamed the parents, the teachers, the classrooms, the resources, the textbooks, the principals, and even the students. Listing all the problems and all the suggested remedies could fill this book many times over.

There are thousands of studies promulgating claims that this method works or that innovation works. We have a rich educational research base, but rarely is it used by teachers, and rarely does it lead to policy changes that affect the nature of teaching. It may be that the research is written in a non-engaging style for teachers, or maybe when research is presented to teachers it is done in a manner that fails to acknowledge that teachers come to research with strong theories of their own about what works (for them). Further, teachers are often very "context specific", as the art for many of them is to modify programs to fit their particular students and teaching methods—and this translation is rarely acknowledged.

How can there be so many published articles, so many reports providing directions, so many professional development sessions advocating this or that method, so many parents and politicians inventing new and better answers, while classrooms are hardly different from 200 years ago (Tyack & Cuban, 1995)? Why does this bounty of research have such little impact? One possible reason is the past difficulties associated with summarizing and comparing all the diverse types of evidence about what works in classrooms. In the 1970s there was a major change in the manner that we reviewed the research literature. This approach offered a way to tame the massive amount of research evidence so that it could offer useful information for teachers. The predominant method had always been to write a synthesis of many published studies in the form of an integrated literature review. However in 1976 Gene Glass introduced the notion of meta-analysis-whereby the effects in each study, where appropriate, are converted to a common measure (an effect size), such that the overall effects could be quantified, interpreted, and compared, and the various moderators of this overall effect could be uncovered and followed up in more detail. Chapter 2 will outline this method in more detail. This method soon became popular and by the mid 1980s more than 100 meta-analyses in education were available. This book is based on a synthesis (a method referred to by some as meta-meta-analysis) of more than 800 meta-analyses about influences on learning that have now been completed, including many recent ones. It will develop a method such that the various innovations in these meta-analyses can be ranked from very positive to very negative effects on student achievement. It demonstrates that the reason teachers can so readily convince each other that they are having success with their particular approach is because the reference point in their arguments is misplaced. Most importantly, it aims to derive some underlying principles about why some innovations are more successful than others in influencing student achievement.

An explanatory story, not a "what works" recipe

The aim is to provide more than a litany of "what works", as too often such lists provide yet another set of recommendations devoid of underlying theory and messages, they tend to not take into account any moderators or the "busy bustling business" of classrooms, and often they appeal to claims about "common sense". If common sense is the litnus test then everything could be claimed to work, and maybe therein lies the problems with teaching As Glass (1987) so eloquently argued when the first What Works: Politics and research was released, such appeals to common sense can mean that there is no need for more research dollars. Such claims can ignore the realities of classroom life, and they too often mistake correlates for causes. Michael Scriven (1971; 1975; 2002) has long written about mistaking correlates of learning with causes. His claim is that various correlates of school outcomes, say the use of advance organizers, the maintenance of eye contact, or high time on task, should not be confused with good teaching. While these may indeed be correlates of learning, it is still the case that good teaching may include none of these attributes. It may be that increasing these behaviors in some teachers also leads to a decline in other attributes (e.g., caring and respect for students). Correlates, therefore, are not to be confused with the causes.

For example, one of the major results presented in this book relates to increasing the amount of feedback because it is an important correlate of student achievement. However, one should not immediately start providing more feedback and then await the magical increases in achievement. As will be seen below, increasing the amount of feedback in order to have a positive effect on student achievement requires a change in the conception of what it means to be a teacher; it is the feedback to the teacher about what students can and cannot do that is more powerful than feedback to the student, and it necessitates a different way of interacting and respecting students (but more on this later). It would be an incorrect interpretation of the power of feedback if a teacher were to encourage students to provide more feedback. As Nuthall (2007) has shown, 80% of feedback a student receives about his or her work in elementary (primary) school is from other students. But 80% of this studentprovided feedback is incorrect! It is important to be concerned about the climate of the classroom before increasing the amount of feedback (to the student or teacher) because it is critical to ensure that "errors" are welcomed, as they are key levers for enhancing learning. It is critical to have appropriately challenging goals as then the amount and directedness of feedback is maximized. Simply applying a recipe (e.g., "providing more feedback") will not work in our busy, multifaceted, culturally invested, and changing classrooms.

The wars as to what counts as evidence for causation are raging as never before. Some have argued that the only legitimate support for causal claims can come from randomized control trials (RCTs, i.e., trials in which subjects are allocated to an experimental or a control group according to a strictly random procedure). There are few such studies among the many outlined in this book, although it could be claimed that there are many "evidence-informed" arguments in this book. While the use of randomized control trials is a powerful method, Scriven (2005) has argued that a higher gold standard relates to studies that are capable of establishing conclusions "beyond reasonable doubt". Throughout this book, many correlates will be presented, as most meta-analyses seek such correlates of enhanced student achievement. A major aim is to weave a story from these data that has some convincing power and some coherence, although there is no claim to make these "beyond reasonable doubt". Providing explanations is sometimes more difficult than identifying causal effects,

Most of these claims about design and RCTs are part of the move towards evidencebased decision making, and the current debate about influences on student learning is dominated by discussion of the need for "evidence". Evidence-based this and that are the buzz words, but while we collect evidence, teachers go on teaching. The history of teaching over the past 200 years has attested the enduring focus of teachers on notions of "what works" despite the number of solutions urging teachers to move in a different direction. Such "what works" notions rarely have high levels of explanatory power. The model I will present in Chapter 3 may well be speculative, but it aims to provide high levels of explanation for the many influences on student achievement as well as offer a platform to compare these influences in a meaningful way. And while I must emphasize that these ideas are clearly speculative, there is both solace and promise in the following quotation from Popper:

Bold ideas, unjustified anticipations, and speculative thought, are our only means for interpreting nature: our only organon, our only instrument, for grasping her. And we must hazard them to win our prize. Those among us who are unwilling to expose their ideas to the hazard of refutation do not take part in the scientific game.

(Popper, K. R., 1968, p. 280)

The challenge 5

□ Students

Parents

Principals

Teachers

While we collect evidence, teachers go on teaching

As already noted, the practice of teaching has changed little over the past century. The "grammar" of schooling, in Tyack and Cuban's (1995) terms, has remained constant: the age-grading of students, division of knowledge into separate subjects, and the self-contained classroom with one teacher. Many innovations have been variously "welcomed, improved, deflected, co-opted, modified, and sabotaged" (p. 7), and schools have developed rules and cultures to control the way people behave when in them. Most of us have been "in school" and thus know what a "real school" is and should be. The grammar of schooling has persisted partly because it enables teachers to discharge their duties in a predictable fashion, cope with the everyday tasks that others expect of them, and provide much predictability to all who encounter schools.

One of the "grammars of schooling" is that students are to be made responsible for their learning. This can easily turn into a conception that some students are deficient in their desire for, and achievements from teaching. As Russell Bishop and his colleagues have demonstrated, such deficit thinking is particularly a problem when teachers are involved with minority students (e.g., Bishop, Berryman, & Richardson, 2002). From their interviews, they illustrated that the influences on Māori students' educational achievement differed for each of parents, students, principals, and teachers (Figure 1.1). Students, parents, and principals see the relationships between teachers and students as having the greatest influence on Māori students' educational achievement. In contrast, teachers identify the main influences on Māori students' educational achievement as being Māori students themselves, their homes and/or the structure of the schools. Teachers engage in the discourse of the child and their home by pathologising Māori students' lived experiences and by explaining their lack of educational



Figure 1.1 Percentage of responses as to the claimed influences on student learning by students, parents, principals, and teachers

achievement in deficit terms. My colleague Alison Jones calls this type of thinking a "discourse of disadvantage" (Jones & Jacka, 1995). They do not see themselves as the agents of influence, see very few solutions, and see very little that they can do to solve the problems.

From their extensive classroom observations, analyses of achievement results, and working with teachers of minority students, Bishop *et al.* have devised a model of teaching Māori students based on caring for all students, and the primacy of the act of teaching. The major features of Bishop's model include the creation of a visible, appropriate context for learning such that the student's culture is involved in a process of co-learning, which involves the negotiation of learning contexts and content. The teacher provides supportive feedback and helps students to learn by acknowledging and using the students' prior knowledge and experiences, and monitoring to check if students know what is being taught, what is to be learnt, or what is to be produced. It involves the teacher teaching the students something, instructing them how to produce something, and giving them instructions as to the processes of learning. This is a high level of teaching activity, indeed.

Concluding comments

This introduction has highlighted the amazing facility of those in the education business to invent solutions and see evidence for their pet theories and for their current actions. Everything seems to work in the improvement of student achievement. There are so many solutions and most have some form of evidence for their continuation. Teachers can thus find some support to justify almost all their actions—even though the variability about what works is enormous. Indeed, we have created a profession based on the principle of "just leave me alone as I have evidence that what I do enhances learning and achievement".

One aim of this book is to develop an explanatory story about the key influences on student learning—it is certainly not to build another "what works" recipe. The major part of this story relates to the power of directed teaching, enhancing what happens next (through feedback and monitoring) to inform the teacher about the success or failure of their teaching, and to provide a method to evaluate the relative efficacy of different influences that teachers use.

It is important from the start to note at least two critical codicils. Of course, there are many outcomes of schooling, such as attitudes, physical outcomes, belongingness, respect, citizenship, and the love of learning. This book focuses on student achievement, and that is a limitation of this review. Second, most of the successful effects come from innovations, and these effects from innovations may *not* be the same as the effects of teachers in regular classrooms—the mere involvement in asking questions about the effectiveness of any innovation may lead to an inflation of the effects. This matter will be discussed in more detail in the concluding chapter, where an attempt is made to identify the effects of "typical" teachers compared to "innovations" in teaching. Indeed, the role of "teaching as intervention" is developed throughout the chapters in this book.

The nature of the evidence A synthesis of meta-analyses

It is the mark of an educated man ... that in every subject he looks for only so much precision as its nature permits.

(Aristotle, 350BC)

This chapter outlines the methodology relating to the evidence used in the remainder of this book. The fundamental unit of analysis is 800+ meta-analyses and how the major results from these studies can be placed along a single continuum. The chapter then outlines some of the problems of meta-analyses, discusses some of the previous attempts to synthesize meta-analyses, and then introduces some of the major overall findings from the synthesis of the 800+ meta-analyses.

Would it not be wonderful if we could create a single continuum of achievement effects, and locate all possible influences of achievement on this continuum? Figure 2.1 shows one possible depiction of this continuum.

Influences on the left of this continuum are those that decrease achievement, and those on the right increase achievement. Those near the zero point have no influence on achievement outcomes.

The next task was to adopt an appropriate scale so that as many outcomes as possible from thousands of studies are converted to this single scale. This was accomplished using effect sizes, and this scale has been among the marvelous advances in the analysis of research studies over the past century. An effect size provides a common expression of the magnitude of study outcomes for many types of outcome variables, such as school achievement. An effect size of d = 1.0 indicates an increase of one standard deviation on the outcome—in this case the outcome is improving school achievement. A one standard deviation increase is typically associated with advancing children's achievement by two to three years, improving the rate of learning by 50%, or a correlation between some variable (e.g., amount of homework) and achievement of approximately r = 0.50. When



Figure 2.1 An achievement continuum

implementing a new program, an effect size of 1.0 would mean that, on average, students receiving that treatment would exceed 84% of students not receiving that treatment.

Cohen (1988) argued that an effect size of d = 1.0 should be regarded as a large, blatantly obvious, and grossly perceptible difference, and as an example he referred to the difference between the average IQ of PhD graduates and high school students. Another example is the difference between a person at 5'3" (160 cm) and 6'0" (183 cm)—which would be a difference visible to the naked eye. The use of effect sizes highlights the importance of the magnitude of differences, which is contrary to the usual emphasis in much of our research literature on statistical significance. Cohen (1990) has commented that "under the sway of the Fisherian scheme [or dependence on statistical significance], there has been little consciousness of how big things are ... science is inevitably about magnitudes ... and meta-analysis makes a welcome force toward the accumulation of knowledge" (pp. 1309–1310).

Thus, we have a continuum and a scale (effect size) to ascertain which of the many possible influences affect achievement. Many textbooks detail how effect sizes can be calculated from various summary statistics such as *t*-tests, ANOVAs, repeated-measures (e.g., Glass, 1977; Glass, McGaw, & Smith, 1981; Hedges & Olkin, 1985). Statistically, an effect size can be calculated in two major ways:

Effect size =
$$[Mean_{treatment} - Mean_{control}]/SE$$

or

Effect size = $[Mean_{end of treatment} - Mean_{beginning of treatment}]/SD$

where SD is the pooled sample standard deviation. There are many minor modifications to these formulas, and for more detail the interested reader is referred to Glass, McGaw, & Smith (1981); Rosenthal (1991); Hedges & Olkin (1985); Hunter & Schmidt (1990); and Lipsey & Wilson (2001).

As an example of synthesizing meta-analyses, take an examination of five meta-analyses on homework: Cooper (1989; 1994); Cooper, Robinson, & Patall (2006); DeBaz (1994); Paschal, Weinstein, & Walberg (1984). Over these five meta-analyses there were 161 studies involving more than 100,000 students, which investigated the effects of homework on students' achievement. The average of all these effect sizes was d = 0.29, which can be used as the best typical effect size of the influence of homework on achievement. Thus, compared to classes without homework, the use of homework was associated with advancing children's achievement by approximately one year, improving the rate of learning by 15%, about 65% of the effects were positive (that is, improved achievement), 35% of the effects were zero or negative, and the average achievement level of students in classes that prescribed homework. However, an effect size of d = 0.29 would not, according to Cohen (1988), be perceptible to the naked eye, and would be approximately equivalent to the difference between the height of a 5'11" (180 cm) and a 6'0" (182 cm) person.

Thus it is possible to devise a unidimensional continuum such as shown in Figure 2.1 that can allow the various effects on achievement to be positioned as they relate to each other. The scale is expressed in effect sizes (or standard deviation units) such that 1.0 is an unlikely—although a very obvious—change in achievement, and 0.0 is no change at all.

This continuum provides the measurement basis to address the question of the relative effects of many factors on achievement.

An alternative way of considering the meaning of an effect size was suggested by McGraw and Wong (1992). They introduced a measure called the common language effect size indicator, which is the probability that a score sampled from one distribution will be greater than a score sampled from some other distribution. Consider as an example the difference in height of the average woman (5'4"/162.5 cm) and the average male (5'10"/177.5 cm), which is a *d* of 2.0. This *d* translates into a common language effect (CLE) of 92 percent. Thus we can estimate that in any random pairing the probability of the male being taller than the female is d = 0.92; or that in 92 out of 100 blind dates the male will be taller than the female. Now, using the example above, consider the d = 0.29 from introducing homework (throughout this book effect sizes are abbreviated, following tradition, to *d*). The CLE is 21 percent, so that in 21 times out of 100, introducing homework into schools will make a positive difference, or 21 percent of students will gain in achievement compared to those not having homework. Or, if you take two classes, the one using homework will be more effective 21 out of a 100 times. In all examples in this book, the CLE is provided to assist in interpreting the effect size.

We do need to be careful about ascribing adjectives such as small, medium, and large to these effect sizes. Cohen (1988), for example, suggested that d = 0.2 was small, d =0.5 medium, and d = 0.8 large, whereas the results in this book could suggest d = 0.2 for small, d = 0.4 for medium, and d = 0.6 for large when judging educational outcomes. In many cases this would probably be reasonable, but there are situations where this would be just too simple. Consider, for example, the effects of an influence such as behavioral objectives, which has an overall small effect of d = 0.20 (see Chapter 9), and reciprocal teaching, which has an overall large effect of d = 0.74. It may be that the cost of implementing behavioral objectives is so small that it is worth using them to gain an influence on achievement, albeit small, whereas it might be too expensive to implement reciprocal teaching to gain the larger effect. Instead of considering only the size of an effect, we should be looking for patterns in the various effect sizes and the causal implications across effect sizes, and making policy decisions on an overall investigation of the differences in effect sizes.

Further, there are many examples that show small effects may be important. A vivid example comes from medicine. Rosenthal and DiMatteo (2001) demonstrated that the effect size of taking low dose aspirin in preventing a heart attack was d = 0.07, indicating that less than one-eighth of one percent of the variance in heart attacks was accounted for by using aspirin. Although the effect size is small, this translates into the conclusion that 34 out of every 1,000 people would be saved from a heart attack if they used low dose aspirin on a regular basis. This sounds worth it to me.

Meyer *et al.* (2001) list other seemingly small effect sizes with important consequences: the impact of chemotherapy on breast cancer survival (d = 0.12), the association between a major league baseball player's batting average and success in obtaining a hit in a particular instance at bat (r = 0.06), the value of antihistamines for reducing sneezes and a runny nose (d = 0.22), and the link between prominent movie critics' reviews and box office success (d = 0.34).

Even more interestingly, it can be possible to identify various moderators that may enhance or detract from the overall average effect. For example, to use the homework case discussed above, it may be that males have greater improvements (i.e., have a higher effect

size) than females, younger students' achievement gains may be different from older ones', the effects may be greater in mathematics than reading. And indeed, the effects do decrease with age: primary students gain least from homework (d = 0.15) and secondary students have greater gains (d = 0.64, see Chapter 10).

Also, the nature of the achievement outcome may turn out to be critical. That is, when one is seeking influences on a very specific, narrow outcome (e.g., improvement in addition, understanding of phonics), then it may be likely that the effect size will be greater than when one is seeking influences on a more generalizable, wider concept (e.g., numeracy or reading achievement). While the synthesis of research on the effects of narrow or wide influences (Hattie, 1992) did not find such differences, it is still important to be aware of the potential of this moderator.

Problems with meta-analysis

Glass (2000) celebrated the 25th anniversary of the invention of the term "meta-analysis" (see also Hunt, 1997) by noting the growth of interest in meta-analysis shifting from an original "preoccupation of a very small group of statisticians" to a current "minor academic industry" (Glass, 2000, p. 1). About 25 percent of all articles in *Psychological Bulletin* have the term "meta-analysis" in the title, and he particularly noted the adoption of the method in medicine. Not surprisingly, given this growth, there remain many criticisms of meta-analysis. A common criticism is that it combines "apples and oranges" and such combining of many seemingly disparate studies is fraught with difficulties. It is the case, however, that in the study of fruit nothing else is sensible. The converse argument is absurd: no two things can be compared unless they are the same! Glass argued that "The question of 'sameness' is not an *a priori* question at all; apart from being a logical impossibility, it is an empirical question" (2000, p. 2). No two studies are the same and the only question of interest is how they vary across the factors we conceive as important.

Another criticism, which Cronbach (1982) referred to as the "flat earth society", is that meta-analysis seeks the *big facts* and often does not explain the complexity nor appropriately seek the moderators. However, meta-analysis indeed can seek moderators, and, as will be seen throughout this book, classrooms are places where complexities abound and all participants constantly try to interpret, engage or disengage, and make meaning out of this variegated landscape. While there are many common themes, sometimes "averages do not do it justice" (Glass, 2000, p. 9). However, the issue (which will be discussed throughout this book) is that the generalizability of the overall effect is an empirical issue, and, as will be seen, there are far fewer moderators than are commonly thought.

A further criticism is that the findings from meta-analysis are based on historical claims—that is, they are based on "past" studies, and the future is not so bound by what worked yesterday. It is critical to always appreciate that the meta-analyses in this book are indeed historical—that is what a research review is: a synthesis of published studies. The degree to which these past studies influence today's or tomorrow's schools is an interpretative issue for the reader.

Eysenck (1984) has been particularly critical of the use of low quality studies in any synthesis, promoting the cliché "garbage in—garbage out". In meta-analysis, it is possible to address this question by ascertaining if the effects are affected by quality, and in general they are not. For example, Lipsey and Wilson (1993) summarized 302 meta-analyses in psychology and education, and used a number of outcomes (besides achievement) in

their analyses (the overall effect was d = 0.50, SD = 0.29). They found no differences between studies that only included random versus non-random design studies (d = 0.46vs. d = 0.41), or between high (d = 0.40) and low (d = 0.37) quality studies. There was a bias upwards from the published (d = 0.53) compared to non-published studies (d = 0.39), although sample size was unrelated to effect size (d = -0.03). Sipe and Curlette (1996) found no relationship between the overall effect size of 97 meta-analyses (d = 0.34) and sample size, number of variables coded, type of research design, and a slight increase for published (d = 0.46) versus unpublished (d = 0.36) meta-analyses. There is one exception that can be predicted from the principles of statistical power: if the effect sizes are close to zero, then the probability of having high confidence in this effect is probably related to the sample size (see Cohen, 1988; 1990).

There is every reason to check the effects of quality, but no reason to throw out studies automatically because of lower quality. An excellent example is the recent synthesis by Torgerson *et al.* (2004), who identified 29 studies from a total of 4,555 potentially relevant papers reporting evaluations of interventions in adult literacy and/or numeracy that were published between 1980 and 2002. Their criterion of acceptance was that only "quality" studies—that is, those studies that used randomized controlled trials—were selected. To decide that it is worthwhile to include only certain types of designs or only studies meeting some criteria of quality presupposes that the studies using only the specified designs or levels of quality are the best representatives of the population estimates. This is speculation, and by using meta-analysis these concerns are subject to verification.

When the studies from Torgerson *et al.* (2004) are examined, it is clear that many of their randomized control studies were of low quality. The median sample size was only 52, and given there were at least two groups (experimental and control) the "typical" study had only 26 people in each group. The average attrition rate was 66 percent, so two-thirds of each sample did not complete the study. It would have been more defensible to include all possible studies, code them for the nature of the experimental design *and* for the quality of the study, and then use meta-analysis techniques to address whether the effects differed as a consequence of design and quality. The aim should be to summarize all possible studies regardless of their design—and then ascertain if quality is a moderator to the final conclusions (see Benseman, Sutton, & Lander, 2005 for a full analysis).

As noted in Chapter 1, Scriven (2005) has argued that a more critical criterion for all scientific conclusions is "beyond reasonable doubt (BRD)", and in some cases randomized studies do not come close to being beyond reasonable doubt. "It seems more appropriate to think of 'gold standard' designs in causal research as those that meet the BRD standard, rather than those that have certain design features ... The existence of more threats to internal or external validity in quasi-experimental designs does not entail a reduction of validity for well-done studies below BRD levels" (pp. 45–46). Scriven noted that one of the advocates of random controlled designs, Cook (2004), claimed that "Interpreting [randomized control trial] results depends on many other things—an unbiased assignment process, adequate statistical power, a consent process that does not distort the populations to which results can be generalized, and the absence of treatment-correlated attrition, resentful demoralization, treatment seepage and other unintended products of comparing treatments. Dealing with these matters requires observation, analysis and argumentation."

critical education questions. Design method and quality of studies are mediators, not prior conditions for choosing studies in a synthesis of studies.

A more statistical concern is that there can be quite a difference depending on whether the author of the meta-analysis used a random or a fixed model to calculate the effect sizes. The fixed effects model can be viewed as a special case of the random model where the variance of the universe effect size is zero; the random model allows generalization to the entire research domain whereas the fixed model allows an estimate of *one* universe effect size underlying all studies available (Kisamore & Brannick, 2008; Schulze, 2004). Typically, but not necessarily, the mean effect size from estimates based on the random model can be appreciably higher than when a fixed model is used. Hence, combining or comparing effects generated from the two models may differ solely because different models are used and not as a function of the topic of interest. Given that the majority of meta-analyses so far published have used the fixed effect model, then this fixed model has been used in this book. Where effects have been based on the random model and this seems to make a difference to the means, this is noted.

Previous attempts at synthesizing meta-analyses

There have been previous attempts to synthesize across meta-analyses. For example, I have published a study based on 134 meta-analyses of studies of educational innovations (Hattie, 1987; 1992). This research concluded that educational innovations can be expected to change average achievement outcomes by 0.4 standard deviations and affective outcomes by 0.2 standard deviations. Some overall findings were drawn about the factors above and below this average benchmark. Innovation, for example, was a theme underlying most of these positive effects. That is, a constant and deliberate attempt to improve the quality of learning on behalf of the system, the principal, and the teacher, typically related to improved achievement. The implementation of innovations probably captures the enthusiasm of the teacher implementing the innovation and the excitement of the students attempting something innovative. Often this has been explained as an experimental artifact in terms of a Hawthorne effect. However, another reason is that when teachers introduce innovation there can be a heightened attention to what is making a difference and what is not, and it is this attention to what is not working that can make the difference—feedback to the teacher about the effects of their actions!

I realized that the most powerful single influence enhancing achievement is feedback. This led me on a long journey to better understand this notion of feedback. After researching and reviewing feedback from a student's perspective (e.g., help-seeking behaviors) and from a teacher to student perspective (e.g., better comments on tests, increasing the amount of feedback in a class), it dawned on me that the most important feature was the creation of situations in classrooms for the teachers to receive more feedback about their teaching—and then the ripple effect back to the student was high (Hattie & Timperley, 2007). Indeed, my team and I have devised a computer-based classroom assessment tool primarily focused on enhancing such feedback (see www.asTTle.org.nz)—but that is another story.

When investigating the continuum of achievement, there is remarkable generality—remarkable because of the preponderance of educational researchers and teachers who argue for treating students individually, and for dealing with curriculum areas as if there were unique teaching methods associated with English, mathematics, and so on. The findings from this synthesis apply, reasonably systematically, to all age groups, all curriculum areas, and to most teachers. It did not seem to matter whether the achievement outcomes were broad or narrow. The average effects of broad constructs and narrow outcomes were slightly lower (d = 0.23) compared with those categorized into broad constructs and broad outcomes (d = 0.43), narrow constructs and narrow outcomes (d = 0.37), and narrow constructs and broad outcomes (d = 0.35). Generality is the norm, but, as with many things, there are exceptions.

The majority of the findings of the meta-analyses were derived from studies conducted in English-speaking, highly developed countries (particularly, but not exclusively, the United States). We should not generalize the findings of these meta-analyses to non-English speaking, or non-highly developed countries. Note, for example, the results of a study by Heyneman and Loxley (1983), drawing on 52,252 elementary school-age pupils, 12,085 teachers, and 2,710 classes from 29 developing countries. They concluded that, relative to high-income countries, academic achievement in low-income countries was affected more by pupils' social status and less by teacher quality.

Kulik and Kulik (1989) reviewed more than 100 meta-analyses, including those relating to instructional methods and design. They concluded that "most of the well-known systems devised for improving instruction have acceptable records in evaluation studies" (p. 289). This was an appropriately cautious claim at that early stage of synthesizing across metaanalyses. They concluded that there were promising effects from curricular innovations (especially in science), and suggested that it was important to be cautious about effects from teacher education programs (which were lower than anticipated). They claimed that large effects were not the norm, although there were few negative effects. The major message for teachers was that there were many advantages with providing clear definitions of learning tasks for students, having a requirement of mastery on class activities and quizzes, and providing increased feedback, but policies relating to reorganizing classrooms did not get much support. These messages of learning intentions, success criteria, direct teaching, and the power of feedback—rather than being concerned with structural adaptations—are still powerful two decades later.

Walberg used my earlier synthesis (Hattie, 1987) to defend his nine-factor "Education Productivity" model, which he argued incorporated the three major psychological causes of learning (Reynolds & Walberg, 1998). The first was student aptitude (prior achievement, d = 0.92; age or maturation, d = 0.51; motivation, self-concept, willingness to persevere on learning tasks d = 0.18). The second was instruction (time in learning, d = 0.47; quality of teaching, d = 0.18). The third was psychological environments (morale or student perceptions of classroom social group, d = 0.47; home environment, d = 0.36; peer group outside school, d = 0.20; minimal leisure-time mass media exposure, particularly television, d = 0.20). More recently he argued that "each of the first five factors—prior achievement, development, motivation, and the quantity and quality of instruction—seems necessary for learning in school. Without at least a small amount of each, the student may learn little ... (each) appears necessary but insufficient by itself for effective learning" (Walberg, 2006, pp. 103–106). Quality is an important enhancement of study time, and the four psychological environments expand and enhance learning time.

Marzano (1998) was critical of these attempts by me, Walberg, and others, claiming that basing a synthesis on "brand names" could be misleading. For example, he argued that the categories we used in our syntheses were too broad and included too many varied treatments, and that instead the categories used should be specific and functional

enough to provide guidance for classroom practice. He used four basic building blocks in his synthesis: knowledge (d = 0.60), the cognitive system (d = 0.75), the meta-cognitive system (d = 0.55), and the self-system (d = 0.74). Marzano used 4,057 effect sizes and found an overall effect size of d = 0.65. (This overall effect is somewhat larger than reported later in this book, as Marzano did not include many of the school and structural influences.) He reported on eight moderators:

- 1 whether the technique was designed for use by the teacher (d = 0.61) or the student (d = 0.73);
- 2 the degree of specificity of the influence (he argued that the more specific the influence, the higher the effect, although the means were d = 0.67, d = 0.64, d = 0.64 for least to most specific);
- 3 grade level of students (no differences);
- 4 student ability (low d = 0.64; middle d = 0.70; and high d = 0.91);
- 5 duration of treatment (shortened programs < 3 weeks d = 0.69 vs, 4 weeks d = 0.52);
- 6 the specificity of dependent measures in the treatment (very specific d = 0.97, appropriate d = 0.91, and very general d = 0.55);
- 7 methodological quality (no difference);
- 8 publication type (published d = 0.72 vs. unpublished d = 0.64).

From his very systematic review he concluded that the "best way to teach organizing ideas—concepts, generalizations, and principles—appears to be to present those constructs in a rather direct fashion" (Marzano, 1998, p. 106) and then have students apply these concepts to new situations. He regarded the meta-cognitive system as the "engine" or primary vehicle for enhancement of the mental processes within the cognitive system and recommended providing students with clear targets of knowledge and skills, and strategies for the processes involved with what they are learning. Marzano, Gaddy, and Dean (2000) outlined an excellent and extremely fascinating set of implications for teachers and the learning processes deriving from these analyses. In a further re-analysis of these effect sizes, Marzano (2000) argued that 80 percent of the variance in achievement could be accounted for by student effects, 7 percent by school effects, and 13 percent by teacher effects. He then used these estimates to evaluate the effects on student achievement of an ineffective, an average, and an exceptional teacher in an ineffective, an average, and an exceptional school respectively. Average schools and average teachers, although he said they did little harm, also did little to influence students' relative position on the distribution of achievement for all students; ineffective teachers, no matter how effective the school, had a negative impact on the standings of all students, whereas students of exceptional teachers, even in ineffective schools, either maintained or increased achievement, many quite substantially."Exceptional performance on the part of teachers not only compensates for average performance at the school level, but even ineffective performance at the school level" (Marzano, 2000, p. 81).

Synthesizing the meta-analyses

This book is not another meta-analysis. There are hundreds of those. Instead, this book aims to synthesize over 800 meta-analyses about the influences on achievement to present a more global perspective on what are and what are not key influences on achievement.

The project started by collecting 134 meta-analyses and proposing a set of common themes as to why some influences were more or less influential than others (Hattie, 1992). Since 1992, this collection of meta-analyses has been supplemented with a large number of other meta-analyses; over the past few years these have all been coded—at the study level—and the current database has a line for each meta-analysis that summarizes and categorizes the study and notes the effect sizes and standard errors which are needed for the calculations reported in this book.

It was possible to locate a total of about 800 meta-analyses, which encompassed 52,637 studies, and provided 146,142 effect sizes about the influence of some program, policy, or innovation on academic achievement in school (early childhood, elementary, high, and tertiary). Topics *not* included are those concerning English as a second language, affective or physical outcomes, and meta-analyses where the number of studies was fewer than four. When the same meta-analysis has been published multiple times, (e.g., when dissertations are rewritten as articles), only the most recent or most accessible is included.

As can be imagined, these effects cover most school subjects (although the majority are reading, mathematics, science, and social studies), all ages, and a myriad of comparisons. These effects are based on many millions of students across the main areas of influence—from the student, the home, the effects of schools, teachers, curricula, and teaching methods and strategies. The total number of students identified in the meta-analyses is large. Only 286 of the meta-analyses included total sample size but together these alone totaled 83 million students. Using the average sample size per study, this would multiply out to about 236 million students in total. However, it is likely that many students would have participated in more than one study, and thus this is a gross estimate of sample size. Even so, it would be safe to conclude that these studies are based on many millions of students.

Appendix A lists all the meta-analyses included in this book, provides the number of studies, people, and effect sizes, along with the average effect size, standard error (if provided), and common language effect. The meta-analyses are listed by the chapters they are referred to in this book. Appendix B lists these influences in their rank order.

The distribution of effect sizes

To start, let us see an overall distribution of all the effect sizes (Figure 2.2) from each of the 800+ meta-analyses. The bars that indicate points on the y-axis represent the number of effects in each category, while the x-axis gives the categories of effect sizes.

There are six immediate implications from Figure 2.2 that are critical to the arguments in this book:

The effects follow a normal distribution. To those immersed in large-scale statistics, this would not be surprising: normality is often, but not necessarily, present when there are large sample sizes. The normal distribution, however, is a consequence of the data and not imposed on it. Given this normal distribution, there are as many influences above the mean effect size as there are below it, and, most importantly, the mean is a reasonably good indicator of all the influences on achievement.

2 Almost everything works. Ninety percent of all effect sizes in education are positive. Of the ten percent that are negative, about half are "expected" (e.g., effects of disruptive students); thus about 95 percent of all things we do have a positive influence on



Figure 2.2 Distribution of effect sizes across all meta-analyses

achievement. When teachers claim that they are having a positive effect on achievement or when a policy improves achievement this is almost a trivial claim: virtually everything works. One only needs a pulse and we can improve achievement.

- 3 Setting the bar at zero is absurd. If we set the bar at zero and then ask that teachers and schools "improve achievement", we have set a very very low bar indeed. No wonder every teacher can claim that they are making a difference; no wonder we can find many answers as to how to enhance achievement; no wonder every child improves. As noted at the outset of this book, it is easy to find programs that make a difference. Raising achievement that is enhancing learning beyond an effect size of d = 0.0 is so low a bar as to be dangerous and is most certainly misleading.
- 4 Set the bar at d = 0.40. The average effect size is d = 0.40. This average summarizes the typical effect of all possible influences in education and should be used as the benchmark to judge effects in education. Effects lower than d = 0.40 can be regarded as in need of more consideration, although (as discussed earlier) it is not as simple as saying that all effects below d = 0.40 are not worth having (it depends on costs, interaction effects, and so on). Certainly effects above d = 0.40 are worth having and a major focus of this book is trying to understand the common denominators of what makes a difference (i.e., the effect sizes above compared to those below d = 0.40). Throughout this book this d = 0.40 effect size is referred to as the hinge-point or h-point, as this is the point on the continuum that provides the hinge or fulcrum around which all other effects are interpreted.
- 5 Innovations are more than teaching: Teachers average an effect of d = 0.20 to d = 0.40 per year on student achievement. This h-point of d = 0.40 does *not* mean that this

is the typical effect of teaching or teachers. It does not mean that merely placing a teacher in front of a class would lead to an improvement of 0.40 standard deviations. In most studies summarized in this book, there is a deliberate attempt to change, improve, plan, modify, and innovate. The best available estimate as to the effects of schooling is based on longitudinal studies. For example, the National Assessment of Educational Progress (NAEP, Johnson and Zwick, 1990) surveyed what students in American schools know and can do in the subject areas of reading, writing, civics, United States history, mathematics, and science. The students were sampled at ages 9, 13, and 17, and the testing has been repeated every two years. The average effect size across the six subject areas was d = 0.24 per year. In our own New Zealand studies, we have estimated that the yearly effect in reading, mathematics, and writing from Years 4 to 13 (N = 83,751) is d = 0.35—although this is not linear: in some years and for some subjects there is more or less growth. The inference for the argument in this book is that teachers typically can attain between d = 0.20 to d = 0.40 growth per year-and that this is to be considered average. They should be seeking greater than d = 0.40 for their achievement gains to be considered above average, and greater than d = 0.60 to be considered excellent.

6 The variance is important. This typical effect size of d = 0.40 may not be uniform across all students or all implementations of any influence. There may be many moderators. For example, the typical effect size of homework is d = 0.29, but the effects are greater for high school students and closer to zero for elementary school students. The major point of this "achievement barometer" or "achievement continuum" is to provide a basis to interpret the effects of change, both the overall effects and effects broken down by important moderators.

The typical effect: the hinge-point

The effect size of 0.40 sets a level where the effects of innovation enhance achievement in such a way that we can notice real-world differences, and this should be a benchmark of such real-world change. It is not a magic number that should become like a p < 0.05cut-off point, but a guideline to begin discussions about what we can aim for if we want to see students change. It provides a "standard" from which to judge effects: it is a comparison based on typical, real-world effects rather than based on the strongest cause possible, or with the weakest cause imaginable. It is not unreasonable to claim that at least half of all implementations, at least half of all students, and at least half of all teachers *can and do* attain this h-point of d = 0.40 change as a consequence of their actions.

An aim of this book is to position the various influences along this continuum, relative to the typical d = 0.40 effect. The fundamental claim is that influences in education are relative: we should judge the success of an innovation relative to d = 0.40 (and certainly not d = 0.0). To return to the homework example used earlier, the typical influence after introducing homework was just below the typical effect across all possible influences. Thus, when the influence of homework is compared to the more usual zero point, those who argue that homework is effective would say "yes", but when the effects from classes without homework are compared to the typical effect across all other influences, then homework is well below an average effect—there are many more innovations that have greater effects. Maybe it is not so surprising that teachers have found that the effect of prescribing homework is not as dramatic as many advocates

and researchers promised. The advocates and researchers compare the outcome to zero, but we should be comparing the effect to alternative innovations. The null hypothesis (d = 0.0) is not the question of interest, so it is no wonder that the answer is misleading; introducing nearly any innovation is better than its absence. The null hypothesis is virtually certain to be false before analysis commences and thus it is uninformative (see Novick & Jackson, 1974).

The main contributors

Table 2.1 presents the average effect for each of the major categories of contributors to learning. The averages of all effects are quite similar with the exception that school differences are far less critical to enhancing achievement: take two students of the same ability and it matters less to which school they go than the influences of the teacher, curricula program, or teaching they experience.

Figure 2.3 presents the number of meta-analyses from these 800+ meta-analyses relative to average d = 0.40 h-point. There are just as many home, student, curricula, and teaching effects above as below the average, more teaching effects above 0.40, and many more school effects below d = 0.40. But averages can hide too much. The remainder of this book works through each of these major categories of influences, chapter by chapter, and aims to more deeply evaluate the underlying causes of what specific innovations and influences are above and below average. Each chapter will work through a number of innovations and influences; sufficient detail is given for each of these innovations to give a sense of the claims, but the primary aim is to draw inferences for the overall model outlined in Chapter 3.

The barometer of influences

We seem to have no barometers of success or failure to show what works and what does not work in education. Yes, we do have tests, lots of them, which we use to evaluate whether students have gained sufficiently. But this is not enough. An influence may "work", but by how much, and how differently from other influences? Some innovations or actions are more influential than others. Instead of asking "What works?" we should be asking "What works best?" as the answers to these two questions are quite different. As has been indicated already, the answer to the first questions is "Almost everything" whereas the answer to the second is more circumscribed—and some things work *better* and some work *worse* relative to the many possible alternatives.

Contribution	No.	Studies	People	Effects	d	SE	CLE
Student	139	11,101	7,513,406	38,282	0.40	0.044	29%
Home	36	2,211	11,672,658	5,182	0.31	0.058	22%
School	101	4,150	4,416,898	13,348	0.23	0.072	16%
Teacher	31	2,225	402,325	5,559	0.49	0.049	35%
Curricula	144	7,102	6,899,428	29,220	0.45	0.076	32%
Teaching	365	25,860	52,128,719	55,143	0.42	0.071	30%
Average	816	52,649	83,033,433	146,626	0.40	0.062	28%

Table 2. Y Average effect for each of the major contributors to learning


Figure 2.3 Number of meta-analyses above and below the h-point

We need a barometer that addresses whether the various teaching methods, school reforms, and so on are worthwhile relative to possible alternatives. We need clear goalposts of excellence for all in our schools to aspire towards, and most importantly, for them to know when they get there. We need a barometer of success that helps teachers to understand which attributes of schooling assist students in attaining these goalposts.

Figure 2.4 outlines one such barometer that has been developed for use throughout this book. The development of this barometer began not by asking whether this or that innovation was working, but whether this teaching worked better than possible alternatives; not by asking whether this innovation was having positive effects compared to not having the innovation, but whether the effects from this innovation were better for students than what they would achieve if they had received alternative innovations.

For each of the many attributes investigated in the chapters in this book, the average of each influence is indexed by an arrow through one of the zones on the barometer. All influences above the h-point (d = 0.40) are labeled in the "Zone of desired effects" as these are the influences that have the greatest impact on student achievement outcomes.



KEY	
Standard error	0.027 (Low)
Rank	88th
Number of meta-analyses	5
Number of studies	161
Number of effects	295
Number of people (4)	105,282

Figure 2.4 A typical barometer of influence

The typical effects from teachers are between d = 0.15 and d = 0.40, as identified from the longitudinal studies discussed above. Any influences in this zone are similar to what teachers can accomplish in a typical year of schooling. The zone between d = 0.0 and d = 0.15 is what students could probably achieve if there was no schooling (and is estimated from the findings in countries with no or limited schooling). Maturation alone can account for much of the enhancement of learning (see Cahan & Davis, 1987). Thus, any effects below d = 0.15 can be considered potentially harmful and probably should not be implemented. The final category includes the reverse effects—those that decrease achievement—and these are certainly not wanted.

The arrow points to the average effect of the various meta-analyses on the particular topic (in the above it is d = 0.29 for the five homework meta-analyses. The variability (or standard error) of the average effect sizes from each meta-analysis is not always easy to determine. Often the information is not provided, and it is well known that the variance is very much related (inversely) to the sample size of studies—the more studies there are, the greater the variance. Across all 800+ meta-analyses the typical standard error of the mean is about d = 0.07—and to provide a broad sense of variance, any influence where the average "spread of effects" is less than d = 0.04 is deemed low, between d = 0.041and d = 0.079 is deemed medium and greater than d = 0.08 is deemed large. While these are crude estimates, it is more important to read the discussion about each influence to ascertain whether important sources of variance could be identified to explain differential effects within that influence. In many cases there is insufficient information to estimate the standard error and thus it is not provided in the summary information. The information under the barometer allows an interpretation of how confident we can be about this summary information: the number of meta-analyses on each category (five in the above case), based on 161 studies, and 295 effect sizes. There were 105,282 students in the four meta-analyses that provided information about sample size. The average effect is d = 0.29, with a standard error of 0.027 (considered "low" relative to all meta-analyses). The effects of homework, in this example, rank 88th of all 138 meta-analyses (see Appendix B).

Relation between effect size and sample size

A funnel plot is often used to examine whether a meta-analysis is based on a biased sample of studies (Light & Pillemer, 1984). The funnel plot is a scatterplot of effect sizes versus the number of studies (in this case), with each data point representing one study. Because meta-analyses with a larger number of studies are more likely to better estimate the effect size, they tend to lie in a narrow band at the top of the scatterplot, while the smaller meta-analyses (with expected greater variation in results) fan out over a larger area at the bottom—thus creating the visual impression of an inverted funnel. As can be seen in Figure 2.5, the results from this synthesis show a reasonably symmetric funnel, indicating a lack of publication bias.

Concluding comments

This chapter sets the scene for the interpretation of the 800+ meta-analyses, which form the fundamental dataset used throughout this book. An achievement continuum has been developed along which the many effects can be located, and the importance of the h-point of d = 0.40 has been emphasized. The barometer of achievement can be used to assist in seeking the explanation of what leads to successful learning that exceeds the d = 0.40 hinge-point.



Figure 2.5 Funnel plot of the effect size and sample size from each meta-analysis

Released

The argument

Visible teaching and visible learning

We think in generalities, but we live in details.

(Whitehead, 1943, p. 26)

This chapter introduces the major findings that will be elaborated in the following chapters. The aim of this book is not to overwhelm with detail from the more than 50,000 studies and 800+ meta-analyses that form the basis of the discussion. Instead, the aim is to build an explanatory story about the influences on student learning and then to convince the reader of the nature and value of the story by working through the evidence to defend it. It is as much theory generation as it is theory appraisal. The art in any synthesis is the overall message, and the simple adage underlying most of the syntheses in this book is "visible teaching and learning". Visible teaching and learning occurs when learning is the explicit goal, when it is appropriately challenging, when the teacher and the student both (in their various ways) seek to ascertain whether and to what degree the challenging goal is attained, when there is deliberate practice aimed at attaining mastery of the goal, when there is feedback given and sought, and when there are active, passionate, and engaging people (teacher, student, peers, and so on) participating in the act of learning. It is teachers seeing learning through the eyes of students, and students seeing teaching as the key to their ongoing learning. The remarkable feature of the evidence is that the biggest effects on student learning occur when teachers become learners of their own teaching, and when students become their own teachers. When students become their own teachers they exhibit the self-regulatory attributes that seem most desirable for learners (self-monitoring, self-evaluation, self-assessment, self-teaching). Thus, it is visible teaching and learning by teachers and students that makes the difference. The following chapters provide the evidence to defend this overall message.

What teachers do matters

The major message is simple—what teachers *do* matters. However, this has become a cliché that masks the fact that the greatest source of variance in our system relates to teachers—they can vary in major ways. The codicil is that what "some" teachers do matters—especially those who teach in a most deliberate and visible manner. When these professionals see learning occurring or not occurring, they intervene in calculated and meaningful ways to alter the direction of learning to attain various shared, specific, and challenging goals. In particular, they provide students with multiple opportunities and alternatives

for developing learning strategies based on the surface *and* deep levels of learning some content or domain matter, leading to students building conceptual understanding of this learning which the students and teachers then use in future learning. Learners can be so different, making it difficult for a teacher to achieve such teaching acts—students can be in different learning places at various times using a multiplicity of unique learning strategies, meeting different and appropriately challenging goals. Learning is a very personal journey for the teacher and the student, although there are remarkable commonalities in this journey for both. It requires much skill for teachers to demonstrate to *all* their students that they can see the students" "perspective, communicate it back to them so that they have valuable feedback to self-assess, feel safe, and learn to understand others and the content with the same interest and concern" (Cornelius-White, 2007, p. 23).

The act of teaching requires deliberate interventions to ensure that there is cognitive change in the student: thus the key ingredients are awareness of the learning intentions, knowing when a student is successful in attaining those intentions, having sufficient understanding of the student's understanding as he or she comes to the task, and knowing enough about the content to provide meaningful and challenging experiences in some sort of progressive development. It involves an experienced teacher who knows a range of learning strategies to provide the student when they seem *not* to understand, to provide direction and re-direction in terms of the content being understood and thus maximize the power of feedback, and having the skill to "get out the way" when learning is progressing towards the success criteria.

Of course, it helps if these learning intentions and success criteria are shared with, committed to, and understood by the learner—because in the right caring and idea-rich environment, the learner can then experiment (be right and wrong) with the content and the thinking about the content, and make connections across ideas. A safe environment for the learner (and for the teacher) is an environment where error is welcomed and fostered—because we learn so much from errors and from the feedback that then accrues from going in the wrong direction or not going sufficiently fluently in the right direction. In the same way, teachers themselves need to be in a safe environment to learn about the success or otherwise of their teaching from others.

To facilitate such an environment, to command a range of learning strategies, and to be cognitively aware of the pedagogical means to enable the student to learn requires dedicated, passionate people. Such teachers need to be aware of which of their teaching strategies are working or not, be prepared to understand and adapt to the learner(s) and their situations, contexts, and prior learning, and need to share the experience of learning in this manner in an open, forthright, and enjoyable way with their students and their colleagues.

We rarely talk about passion in education, as if doing so makes the work of teachers seem less serious, more emotional than cognitive, somewhat biased or of lesser import. When we do consider passion, we typically constrain such expressions of joy and involvement to matters unrelated to our teaching (Neumann, 2006). The key components of passion for the teacher and for the learner appear to be the sheer thrill of being a learner or teacher, the absorption that accompanies the process of teaching and learning, the sensations in being involved in the activity of teaching and learning, and the willingness to be involved in deliberate practice to attain understanding. Passion reflects the thrills as well as the frustrations of learning—it can be infectious, it can be taught, it can be modeled, and it can be learnt. It is among the most prized outcomes of schooling and, while rarely studied

in any of the studies reviewed in this book, it infuses many of the influences that make the difference to the outcomes. It requires more than content knowledge, acts of skilled teaching, or engaged students to make the difference (although these help). It requires a love of the content, an ethical caring stance to wish to imbue others with a liking or even love of the discipline being taught, and a demonstration that the teacher is not only teaching but learning—typically about the students' processes and outcomes of learning.

Learning is not always pleasurable and easy; it requires over-learning at certain points, spiraling up and down the knowledge continuum, and building a working relationship with others in grappling with challenging tasks. This is the power of deliberative practice. It also requires a commitment to seeking further challenges—and herein lies a major link between challenge and feedback, two of the essential ingredients of learning. The greater the challenge, the higher the probability that one seeks and needs feedback, but the more important it is that there is a teacher to provide feedback and to ensure that the learner is on the right path to successfully meet the challenges.

The key to many of the influences above the d = 0.40 h-point is that they are deliberate interventions aimed at enhancing teaching and learning. But the message is to not merely *innovate*—but for us to learn from what makes the difference when teachers innovate. When we innovate we are more aware of what is working and what is not, we are looking for contrary evidence, we are keen to discover any intended and unintended consequences, and we have a heightened awareness of the effects of the innovations on outcomes. In these situations teachers become the learners about their own effects! In any innovation there is deliberate attention to implementation and its effects, there is a degree of challenge, and a valuing of feedback. It is critical that teachers learn about the success or otherwise of their interventions: those teachers who are students of their own effects are the teachers who are the most influential in raising students' achievement. Seeking positive effects on student learning (say, d > 0.40) should be a constant theme and challenge for teachers. As this does not accrue by serendipity or accident, then the excellent teacher must be vigilant to what is working and what is *not* working in the classroom.

A concept of excellent teaching

A story serves to illustrate my claims about excellent teaching. Some time ago one of my Master's students completed a meta-analysis on the effects of various programs on selfconcept of children and adults (Clinton, 1987). The most successful programs were the Outward Bound or Adventure programs. There are four major features of these programs that led to their positive influence. First, Outward Bound programs have an emphasis on the immediate quality of the experience, as well as aiming to have these immediate experiences have an effect on later experiences. That is, there is a planned and intentional transfer of experiences, knowledge, and decisions during the earlier learning experiences to later experiences (see Hattie, Marsh, Neill, & Richards, 1997 for more details). Second, Outward Bound programs set difficult and specific goals that are known to the learner, and then tasks are structured so that participants can attain these goals. The program provides challenging and specific goals (e.g., successfully negotiating a 60-foot cliff by abseiling or rappelling) and then structures situations (e.g., adequate preparation, social support, etc.) so that participants share a commitment to reaching these goals. Third, the program increases the amount and quality of feedback, which is vital to the learning process. The often dangerous and risky situations demand feedback, and the learning intentions and

success criteria are crystal clear. A major function of challenging and specific goals is that they direct attention and effort, and thus the learner is more aware and keen for feedback related to attaining these goals. Fourth, in the Outward Bound program, the instructor is keenly aware of the need to understand and, if necessary, reassess and redirect an individual's coping strategies. Such coping strategies can be cognitive (learning strategies), personal (building of self-efficacy, perseverance in the face of challenge), and social (help seeking, cooperative learning). These four major features are the keys of successful teaching and learning.

As another example, take my involvement in a Bush Search and Rescue Squad which involved teaching the skills and fun of cliff rescues. Consider the following: I am going to take you to the top of a three-storey building and teach you to rappel down the outside of this building. Typically, I would then demonstrate to you how to put on a safety harness, tie the rope in a bowline, and then show you how to lean backwards to commence the descent. In line with the principles of good teaching, I would then ask you the student, to implement this learning. Typically, such a learning situation leads to much care by the students, an enhanced level of interest in what peers are doing, and high levels of helpseeking behaviors to ensure the knowledge of rope-work is correct and harnesses are correctly positioned. The goals are challenging, specific, and visible, and the learners are committed to them! The learning is actively visible and there are high levels of feedback and monitoring. The learner typically "seeks" the feedback. When a novice first gets to the edge, there is a remarkably high level of peer teaching and learning; it is not natural to fall backwards when descending as it is more typical for the feet to precede the head. When finally the student reaches the bottom there is a surge of excitement appreciating that the challenging goal has been reached (it is abundantly clear what the success criteria are!), the experience was exhilarating, and the learning absorbed in the experience itself. Most want to repeat the experience and continue to enjoy meeting the challenging goals. Moreover, all these acts and most of the "thinking" about the task are visible to the teacher and to the learner. This is the heart of the model of successful teaching and learning advocated in this book.

Visible teaching

It is critical that the teaching and the learning are visible. There is no deep secret called "teaching and learning": teaching and learning are visible in the classrooms of the successful teachers and students, teaching and learning are visible in the passion displayed by the teacher and learner when successful learning and teaching occurs, and teaching and learning requires much skill and knowledge by both teacher and student. The teacher must know when learning is correct or incorrect; learn when to experiment and learn from the experience; learn to monitor, seek and give feedback; and know to try alternative learning strategies when others do not work. What is most important is that teaching is visible to the student, and that the learning is visible to the teacher. The more the student becomes the teacher and the more the teacher becomes the learner, then the more successful are the outcomes.

This explanation of visible teaching relates to teachers as activators, as deliberate change agents, and as directors of learning. This does not mean that they are didactic, spend 80 percent or more of the day talking, and aim to get through the curriculum or lesson come what may. Effective teaching is not the drilling and trilling to the less than willing.

When I reviewed the videotapes of many of the best teachers in the United States (via the National Board for Professional Teaching Standards video assessment task) it was stunning how active and involved the best teachers were in the classrooms—it was clear who was in control in those classrooms. The activity was visible and "in the air"; passive was not a word in the vocabulary of these accomplished teachers—learning was not always loud and heated but it was rarely silent and deadening, and it was often intense, buzzing, and risky.

The model of visible teaching and learning combines, rather than contrasts, teachercentered teaching and student-centered learning and knowing. Too often these methods are expressed as direct teaching versus constructivist teaching (and then direct teaching is portrayed as bad, while constructivist teaching is considered to be good). Constructivism too often is seen in terms of student-centered inquiry learning, problem-based learning, and task-based learning, and common jargon words include "authentic", "discovery", and "intrinsically motivated learning". The role of the constructivist teacher is claimed to be more of facilitation to provide opportunities for individual students to acquire knowledge and construct meaning through their own activities, and through discussion, reflection and the sharing of ideas with other learners with minimal corrective intervention (Cambourne, 2003; Daniels, 2001; Selley, 1999; von Glasersfeld, 1995). These kinds of statements are almost directly opposite to the successful recipe for teaching and learning as will be developed in the following chapters.

A model of learning

The major point here is that constructivism is not a theory of teaching, but a theory of knowing and knowledge, and it is important to understand the role of building constructions of understanding. Bereiter (2002) used Popper's three worlds to make sense of much of what we strive for in school: the physical world, the subjective or mental world, and the world of ideas. These three worlds have major parallels with the three worlds of achievement: surface knowledge of the physical world, the thinking strategies and deeper understanding of the subjective world, and the ways in which students construct knowledge and reality for themselves as a consequence of this surface and deep knowing and understanding. This third world, often forgotten in the passion for teaching facts and thinking skills, is entirely created by humans, is fallible but capable of being improved, and can take on a life of its own. Students often come to lessons with already constructed realities (third worlds), which, if we as teachers do not understand them before we start to teach, can become the stumbling blocks for future learning. If we are successful, then the students' constructed realities (based on their surface and deep knowing) and keenness to explore these worlds are the major legacy of teaching. The contents of this third world are not concrete like books, statues, and teapots (see Bereiter, 2002, pp. 62–63): they are more conceptual. It is certainly the case, as Bereiter documents, that "much of what is meant by the shift from an industrial to a knowledge society is that increasing amounts of work are being done on conceptual objects rather than on the physical objects to which they are related" (Bereiter, 2002, p. 65).

The distinctions are not clean cut, as at all three levels we often learn in a haphazard manner. So much teaching is aimed at the first world—the world of ideas and knowledge, and there is also much discussion about the importance of deep knowledge and thinking skills (the second world). But the task of teaching and learning best comes together when we attend to all three levels: ideas, thinking, and constructing.

In many situations, it will be impossible to make a clear distinction between knowing or thinking about the conceptual artifact and knowing or thinking about the material world the conceptual artifact applies to [...] What matters is that we recognize conceptual artifacts as real things, recognize creating and improving them as real work, and recognize understanding them as real understanding.

(Bereiter, 2002, p. 67).

The real work of schooling is to create or add value to conceptual artifacts in the same way that builders add value to building artifacts. It is a world of conjectures, explanations, proofs, arguments, and evaluations.

Similarly, there can be many cultural artifacts particular to a culture and an important aspect of education is to teach these artifacts. For example, in New Zealand Māori culture, there is much importance attached to whānau (family), history, and cultural norms. A major focus of schooling is to therefore enable learners to adopt these cultural artifacts as a key part of their own conceptual artifacts—a way to see their world in a similar manner to how the culture has learnt to see its world, and communicate its worldviews and values. The importance is that this three-level view of achievement allows relations between theory and observation, personal and cultural belief and observation, and between personal belief and theory.

... there are the relations between different theories, different phenomena, and different people's readings of the same phenomena. None of these relations are easy. They are all inferential and highly problematic. But they are what people work on when they are building scientific knowledge.

(Bereiter, 2002, p. 91)

Bereiter claimed that there are a number of commitments to knowledge improvement: the third world is not limited to accepted, verified, or important knowledge objects. It can include discredited theories, crank notions, unsolved problems, and new ideas that may or may not gather a following. In this respect, "the third world is more inclusive than the canons of liberal education. This inclusiveness goes a long way toward eliminating the split between established knowledge and students' constructive efforts because it places the ideas created by students in the same world as the ideas handed from authoritative sources" (p. 237). "Knowing one's way around in the world of conceptual artifacts affords a wealth of possibilities not open to people who know that world only from a distance, if at all." (p. 238). Knowledge building is an activity directed toward the third world. It is doing something to a conceptual artifact. Bereiter claimed that knowledge building includes thinking of alternatives, thinking of criticisms, proposing experimental tests, deriving one object from another, proposing a problem, proposing a solution, and criticizing the solution. It is more than knowing, mistakenly believing, or doubting some knowledge object.

Educating is more than teaching people to think—it is also teaching people things that are worth learning. Good teaching involves constructing explanations, criticizing, drawing out inferences, finding applications, and there "should never be a need for the teacher to think of ways to inject more thinking into the curriculum. That would be like trying to inject more aerobic exercise into the lives of Sherpa porters" (Bereiter, 2002, p. 380). If the students are not doing enough thinking, something is seriously wrong with the instruction. "If the only justification for an activity is that it is supposed to encourage or improve

thinking, drop it and replace it with an activity that advances students' understanding that increases their mastery of a useful tool" (Bereiter, 2002, p. 381).

Surface, deep, and constructed understanding

There needs to be a major shift, therefore, from an over reliance on surface information (the first world) and a misplaced assumption that the goal of education is deep understanding or development of thinking skills (the second world), towards a balance of surface and deep learning leading to students more successfully constructing defensible theories of knowing and reality (the third world).

For many students, success at school relates to adopting a surface approach to understanding both how and what they should learn, whereas many teachers claim that the goal of their teaching is enhancing deep learning (Biggs & Collis, 1982). Brown (2002), for example, investigated the beliefs about learning of more than 700 15-year-old New Zealand students and 71 of their teachers of English, mathematics, and science. Students argued that learning for them primarily meant exhibiting surface knowledge involving the reproduction of taught material in order to maximize achievement in assessments. In contrast, teachers of these same students claimed that they were teaching towards deep learning outcomes. The students were more governed by the tasks and examinations set by teachers and schools, so, despite claims by teachers, the students were very strategic in concentrating on acquiring sufficient surface and whatever deeper understanding was needed to complete assignments and examinations. (The same phenomenon is especially evident when comparing conceptions of learning by academics and their students; Purdie, 2001.)

Students can be strategic in their approach because most questions and examinations (verbal and written) relate to surface knowledge (Marzano, 1991). For example, Gall (1970) claimed that 60 percent of teachers' questions required factual recall, 20 percent were procedural, and only 20 percent required thought by the students. Other studies have found the proportion of surface thinking questions can be in the order of 80 percent or more (Airasian, 1991; Barnette, Walsh, Orletsky, & Sattes, 1995; Gall, 1984; Kloss, 1988). Teachers' questioning may not elicit deep thinking from students because students understand that questioning is how teachers lead and control classroom activity; in other words, students know that the teacher already knows the answer to the questions (Gipps, 1994; Torrance & Pryor, 1998; Wade & Moje, 2000). So much of daily classroom life is "knowledge telling", and thus surface knowledge is sufficient. Students soon learn that studying on learning with surface strategies or methods (i.e., revision, re-reading, and reviewing of the year's work) leads to success. In contrast, teachers claim to prefer a deep view of learning, usually focused on academic and cognitive development, while at the same time they emphasize surface methods of teaching, usually with the defense that this is what is required in order to prepare students for high-stakes qualification examinations or assessments. This emphasis on surface approaches means that students tend to experience very few opportunities or demands for deep thinking in contemporary classrooms.

To be more specific, *surface* learning involves a knowing or understanding of ideas or facts. In contrast, the two *deep* processes—relational and elaborative—constitute a change in the quality of thinking that is cognitively more challenging than surface questions. Relational responses require integration of at least two separate pieces of given knowledge, information, facts, or ideas. In other words, relational questions require learners to impose

an organizing pattern on the given material. Elaborative or extended abstract responses require the learner to go beyond the given information, knowledge, or ideas, and deduce a more general rule or proof that applies to all cases. In such cases, the learner is forced to think beyond the given and bring in related, prior, or new knowledge, ideas, or information in order to create an answer, prediction, or hypothesis that extends to a wider range of situations. From these surface and deep knowing and understandings the learner can construct notions or ideas that then shape the ways they engage in surface and deep learning (the third world of constructed understanding).

These three types of understanding—surface, deep, and constructed or conceptual understanding—are built on the Biggs and Collis (Biggs & Collis, 1982; Collis & Biggs, 1979) SOLO model of student learning that has proven most valuable both in developing models of teaching and learning and also in our understanding of assessment (Hattie & Purdie, 1998; Hattie & Brown, 2004). These forms of building on surface knowledge to develop deep knowledge are becoming common in the research on educational psychology and assessment. It is intriguing to note that the major revision of Bloom's Taxonomy (Anderson, Krathwohl, & Bloom, 2001) introduced four similar levels; factual knowledge (how to be acquainted with a discipline or solve problems in it); conceptual knowledge (interrelationships among elements within a large structure that enable them to function together); procedural knowledge (how to do something, methods of inquiry); and meta-cognitive knowledge (knowledge of cognition in general as well as awareness and knowledge of one's own cognition). This is a major advance on the better-known Bloom's Taxonomy, which confuses levels of knowing with forms of knowledge (see Hattie & Purdie, 1998).

It is critical to note that the claim is not that surface knowledge is necessarily bad and that deep knowledge is essentially good. Instead, the claim is that it is important to have the right balance: you need surface to have deep; and you need to have surface and deep knowledge *and* understanding in a context or set of domain knowledge. The process of learning is a journey from ideas to understanding to constructing and onwards. It is a journey of learning, unlearning, and overlearning. When students can move from idea to ideas and then relate and elaborate on them we have learning—and when they can regulate or monitor this journey then they are teachers of their own learning. Regulation, or meta-cognition, refers to knowledge about one's own cognitive processes (knowledge) and the monitoring of these processes (skillfulness). It is the development of such skillfulness that is an aim of many learning tasks, and developing them so there is a sense of self-regulation.

A reminder about outcomes

As noted in the earlier chapters, the focus of this book is on achievement outcomes. Now this notion has been expanded to achievement outcomes across the three worlds of understanding. It may seem intuitively obvious that the influences on learning that aim for surface learning tend to favor more directed, specific goals, whereas those that aim for deep learning tend to favor more inquiry methods. Not so—this is too simple and can be misleading. Sometimes the deeper concepts need more specific and direct teaching, and sometimes the more surface concepts can be learned via inquiry or problem solving.

A major aim is to develop "over-learning" or fluency of achievement. For example, most of us "over-learnt" learning to walk—we forget the trial and error and pain that was

involved when we first learnt to walk; but we can most certainly recognize that struggle to learn to walk when we have a major accident, and must learn this skill anew. We want a sense of fluency and over-learning of worthwhile activities as a major outcome of schooling. There is over-learning when we consider a person fluent in a language or with a musical instrument, or when we consider a student fluent in math, reading, or science. A sufficient level of fluency can lead to other desirable outcomes such as retention, endurance, stability, and application within a domain (Doughty, Chase, & O'Shields, 2004).

When a student attains a high degree of fluency on a topic, then they have more cognitive resources to devote to the next phase in learning. When tasks are very complex for the student, the quality of meta-cognitive skills rather than intellectual ability is the main determinant of learning outcomes (Veenman, Prins, & Elshout, 2002) "because learners have to improvise and use heuristics rather than call upon knowledge and skill components that are associated with intellectual ability" (Prins, Veenman, & Elshout, 2006, p. 377). The novice is more likely to use trial and error strategies, whereas the student with greater knowledge is more likely to search for all possible strategies that might work (Klahr, 2000). The novice aims to produce data, whereas the expert is more interested in data interpretations. The data gathering precedes the data interpretation. These claims are the case for both the learner and for the teacher.

Our cognitive architecture has limitations: we can only remember so many things at once; we can only devote so much cognitive processing power to learning and resolving dilemmas. We build higher order notions or schema to help us retain more in memory at any one time, and we learn various strategies to assist in the learning process. These limitations relate to the notion of cognitive load (e.g., Sweller, 2006). Certainly when first learning new material and ideas we need effective learning strategies and as much cognitive processing power as we can muster. Experts, compared to non-experts, have deeper and more principled problem representations that allow more retrieval and resolution, thus demonstrating how they can effectively use the load on their cognition (e.g., when playing chess, solving equations, reading history). A key difference, however, between experts and novices, is that it is deliberative practice rather than experience that matters—that is, extensive engagement in relevant practice activities for improving performance (as when swimmers swim lap after lap aiming to over learn the key aspects of their strokes, turns, and breathing).

Such deliberative practice activities:

are at an appropriate, challenging level of difficulty, and enable successive refinement by allowing for repetition, given room to make and correct errors, and providing informative feedback to the learning [...] Given that the deliberate practice requires students to stretch themselves to a higher level of performance, it requires full concentration and is effortful to maintain for extended periods.

(van Gog, Ericsson, Rikers, & Paas, 2005, p. 75)

All this practice leads to higher levels of conscious monitoring and control that leads to more refinement, and more higher order understandings of the surface and deeper level notions (Charness, *et al.* 2005). It is not deliberative practice for the sake of repetitive training, but deliberative practice focused on improving particular aspects of the target performance, to better understand how to monitor, self-regulate and evaluate their performance, and reduce errors.

The six factors

The next seven chapters of this book are structured around six topics—an assessment of the respective contributions to achievement from:

- 1 the child;
- 2 the home;
- 3 the school;
- 4 the curricula;
- 5 the teacher;
- 6 the approaches to teaching (two chapters).

Of course, there are likely to be interactions between these (another topic too rarely subjected to study and meta-analysis) and this will be returned to in the final chapter. There may also be moderators of these influences, although remarkably, such moderators are few and far between. What works best appears to be similar across subject, age, and context.

The child

The contributions the child brings to his or her learning include:

- prior knowledge of learning;
- expectations;
- degree of openness to experiences;
- emerging beliefs about the value and worth to them from investing in learning;
- engagement;
- ability to build a sense of self from engagement in learning, and a reputation as a learner.

The child brings prior knowledge of learning to their classroom—from preschool, from their culture, from television, from home, and from the previous year. Much of this prior knowledge leads to expectations by students and teachers about learning. A child is born into and grows up in a world of expectations. These expectations are powerful enhancers of—or inhibitors to—the opportunities provided in schools. They come from the parents, from the family, from siblings, from peers, from schools, from teachers, from media, and from themselves. Their own expectations can be formed powerfully from experiences in classrooms. By the age of eight, so many students have worked out their place in the rankings of the achievement equation. It is therefore a concern that one of the greatest influences on student achievement identified in this book is that of self-reported grades students are very adept at knowing how to rate their performance. If these ratings are too low, then such expectations of performance can set limits of what students see as attainable. Hence, there is power in teachers setting more challenging goals, engaging students in the learning towards these goals, and giving students the confidence to set and attain their goals. A student's own predictions of their performance should not be the barriers to exceeding them, as they are for too many students.

A major way these expectations are manifested in the learning situation is via the

student's dispositions. The key dispositions are the way the student becomes open to experiences, their emerging beliefs about the value and worth to them from investing in learning, and the manner in which they learn that they can build a sense of self from their engagement in the learning enterprise. While these can be changed within schools, they also can be formed and changed in the home, in the playground, via interactions with non-school activities (books, television), and, powerfully, by peers. There are many opportunities for parents and educators to mould dispositions that aid rather than hinder learning, such as developing the child's willingness to engage in learning, the degree that a child aims to enhance his or her reputation that can be gained from being engaged in learning, helping the child attribute success to factors such as effort rather than ability, and developing in the child a positive attitude towards learning. These positive attitudes of openness to experience, willingness to invest in learning, and intellectual engagement can be fostered in preschools, and then developed to a particularly high level in our schools-providing we can ensure that tasks are appropriately challenging to students, and that success is attributed to their investment in the tasks. This can then lead to a sense of reputation enhancement-students derive a sense of self and reputation among peers that they are "learners" (Carroll, Hattie, Durkin & Houghton, 2001). Therein lies success. Such personal dispositions can have a marked impact on the outcomes of schooling.

As will be shown in Chapter 9, having and sharing challenging goals/learning intentions with students is a major condition of successful learning, but on top of this it helps considerably if students share a commitment and sense of engagement to these goals. Many meta-analyses of the effects of intention on behavior have shown that intentions accounted for 28 percent of the variance of behavior, and is highest when students possess actual control over the behavior (e.g., Armitage & Conner, 2001, d = 0.24; Hausenblas, Carron, & Mack, 1997, d = 0.23; Milne, Sheeran, & Orbell, 2000, d = 0.20; Sheeran, 2002, d = 0.27; Webb & Sheeran, 2006, d = 0.29 between intentions and behavior change). Working towards appropriately challenging goals requires many attributes, such as commitment, engagement, openness to experience, and a desire to seek a reputation among peers as a learner. Levin (1988) has often argued that one of the most powerful predictors of health, wealth, and happiness in adult life relates more to the number of years in schooling than to achievement. Hence, a major goal of schools should be to *turn us on* to learning (irrespective of where we fall on the achievement ladder) and to assist us to be open to new experiences in learning.

There are many ways to entice engagement in learning. Steinberg, Brown and Dornbusch (1997) have argued that no manner of school reform will be successful until we first face and resolve the engagement problem—and they note that this is not merely an educational problem, but is "a more general barometer of adolescent malaise" (Steinberg, Brown & Dornbusch, 1997, p. 63). Too many students are "physically present but psychologically absent" (p. 67). Part of the problem is that students can be confused (cannot keep up, or classes are too difficult), also so many are bored (too easy, too few consequences of the learning). When one adds Nuthall's (2005) finding that most of the material taught in a class is already known by the students, and Yair's (2000) claim that students spend 85 percent of their time listening (or pretending to listen) to a teacher talking, then we make it difficult to foster engagement (see also Sirotnik, 1985). We need, claims Steinberg, better indicators of success, more challenging material, higher expectations, and more successful ways to orient students to succeed in school rather than merely helping students avoid the negative consequences of failing to graduate.

The home

Influences from the home on student learning include:

- parental expectations and aspirations for their child;
- parental knowledge of the language of schooling.

The home can be a nurturing place for the achievement of students, or it can be a toxic mix of harm and neglect with respect to enhancing learning. Many parents, however, begin with positive expectations for their children, and these expectations can be critical to the success of children when they go to school. A major concern, however, is the extent to which parents know how to "speak the language of schooling" and thus can advantage their children during the school years; some do not know this language and this can be a major barrier to the home contributing to achievement and to the realization of parents' expectations for their children (Clinton, Hattie, & Dixon, 2007). Schools have an important role in helping parents to learn the language of schooling so that the parents can provide every possible assistance to their children in terms of developing the child's learning and love of learning, and in creating the highest possible shared expectations for learning.

The school

School effects include:

- the climate of the classroom, such as welcoming errors, and providing a safe, caring environment;
- peer influences.

The effects of schools too often are overplayed—particularly in developed countries. Take two students of similar ability; in many developed countries it matters not which school they attend. Many of the school effects are structural (e.g., architecture of school, timetabling differences) or working conditions (e.g., class size; tracking, or streaming, of classes; school finances). Of course these are important, but they do not define the differences in student achievement: they are among the least beneficial influences on student achievement. That has not stopped these structural and working conditions becoming the most discussed issues in education.

Indeed, one of the fascinating discoveries throughout my research for this book is discovering that many of the most debated issues are the ones with the least effects. It is a powerful question to ask why such issues as class size, tracking, retention (that is, holding a student back a grade), school choice, summer schools, and school uniforms command such heated discussion and strong claims. Such cosmetic or "coat of paint" reforms are too common. So many structural claims involve the parents (more homework), lead to more rules (and therefore more rule breakers), have hints of cultural imperatives (quietness and conformity is desired), and often include appeals to common sense (reducing class size is obviously a good thing!). However, the most powerful effects of the school relate to features *within* schools, such as the climate of the classroom, peer influences, and the lack of disruptive students in the classroom—all of which allow students and teachers to make errors and develop reputations as learners, and which provide an invitation to learn.

Purkey (Novak, & Purkey, 2001; Purkey, 1992) has built a theory, known as "Invitational Learning", which works from the meaning of invitational as "offering something beneficial for consideration". His claim is that we need to create schools that invite, or cordially summon students to be involved in the learning process. The model is based on four propositions:

- 1 trust, in that we need to convince not coerce others to engage in what we would like them to consider worthwhile activities;
- 2 respect, in that we adopt caring and appropriate behaviors when treating others;
- 3 optimism, in seeking the untapped potential and uniqueness in others;
- 4 intentionality, in which we create programs by people designed to invite learning.

This is not "niceness" at work, but an approach that places much reliance on the teachers and schools to make learning exciting, engaging, and enduring. Where there are school differences, it is these types of effects that are the most powerful.

The teacher

The teacher contributions to student learning include:

- the quality of teaching—as perceived by the students;
- teacher expectations;
- teachers' conceptions of teaching, learning, assessment, and the students—this relates to teachers' views on whether all students can progress and whether achievement for all is changeable (or fixed), and on whether progress is understood and articulated by teachers;
- teacher openness—whether teachers are prepared to be surprised;
- classroom climate—having a warm socio-emotional climate in the classroom where errors are not only tolerated but welcomed;
- a focus on teacher clarity in articulating success criteria and achievements;
- the fostering of effort;
- the engagement of all students.

The current mantra is that *teachers make the difference*. As noted above, this message, like most simple solutions, is not quite right—it is some teachers undertaking certain teaching acts with appropriately challenging curricula and showing students how to *think* or *strate-gize* about the curricula. Not all teachers are effective, not all teachers are experts, and not all teachers have powerful effects on students. The important consideration is the extent to which they do have an influence on student achievements, and what it is that makes the most difference.

A most critical aspect contributed by the teacher is the quality of their teaching as perceived by the students. Irving (2004) created a student evaluation of high school mathematics teachers based on the National Board for Professional Standards for this domain (www.nbpts.org). After completing a study on the psychometrics of the instrument in New Zealand, he then located a cohort of American teachers who had passed National Board Certification in high school mathematics, and a comparable group who had not passed. He administered student evaluations to both groups. The students were accurate

judges of excellence, and could discriminate between teachers who were experienced and expert from those who were experienced and non-expert. The dimensions that contributed most to this discrimination had a focus on cognitive engagement with the content of the mathematics curriculum, and the development of a mathematical way of thinking and reasoning. It is what teachers get the students to do in the class that emerged as the strongest component of the accomplished teachers' repertoire, rather than what the teacher, specifically, does. Students must be actively involved in their learning, with a focus on multiple paths to problem solving. As mathematical thinkers and problem solvers, the students are also encouraged to go beyond the successful solution of the problem to include the interpretation and analysis of the solution. All the while, students are encouraged to greatly value mathematics and the work that they do in mathematics, and always check the quality of their work to strive for the very best standards. As Irving argued, we should not overlook those who are arguably in the best position to evaluate the teachers-the students who share the classroom with the teacher day in and day out. The myths that students are capricious, and that they are likely to award their teachers high grades was not supported by this research (Irving, 2004). High ratings were not awarded lightly (Bendig, 1952; Tagomori & Bishop, 1995).

There is quite a jump down in the size of the effects to the next contributions related to the teacher: their expectations and their conceptions of teaching. As children are born into a world of expectations, similarly they walk into classrooms with their own expectations to confront teachers who also have expectations of them. Teachers also walk into classrooms with conceptions of teaching, learning, assessment, and the students. We need to better understand these conceptions as it seems they are powerful moderators of the success of these teachers. Having low expectations of the students' success is a self-fulfilling prophecy and it appears that expectations are less mediated by between-student attributes (e.g., gender, race) but held for all students in the teacher's class (Rubie-Davies, 2006, 2007; Rubie-Davies, Hattie, & Hamilton, 2006; Weinstein, 2002). What matters are conceptions of teaching, learning, assessment, and teachers having expectations that *all* students can progress, that achievement for *all* is changeable (and not fixed), and that progress for *all* is understood and articulated. It is teachers who are open to experience, learn from errors, seek and learn from feedback from students, and who foster effort, clarity, and engagement in learning.

The curriculum

Aspects relating to the curriculum that have an influence on student learning include:

- developing a curriculum that aims for the best balance of surface and deep understanding;
- ensuring a focus on developing learning strategies to construct meaning;
- having strategies that are planned, deliberate, and having explicit and active programs that teach specific skills and deeper understanding.

It appears from the many studies reviewed in the subsequent chapters that the major influences on achievement cross curriculum boundaries—the more important attribute is the balance of surface or deep understanding within each curriculum subject, which leads to conceptual clarity. The facility to develop a series of learning strategies for assisting

students to construct meaning from text, develop understanding from numbers, and learn principles is important. These strategies must be planned, deliberate, and explicit, and there need to be active programs to teach specific skills and deeper understanding in the subject areas. Such strategies can then lead to further engagement in the curriculum, leading to the development of problem solving skills, and to the enjoyment of some control over one's learning. This then leads to further developing learning strategies to master content and understanding. A key feature is that many of these strategies can only be enhanced within a domain of knowledge and there can be little transfer (Hattie, Biggs, & Purdie, 1996). This is particularly the case when learning deeper understanding and developing conceptual understanding.

Teaching approaches

Aspects of teaching approaches that are associated with student learning include:

- paying deliberate attention to learning intentions and success criteria;
- setting challenging tasks;
- providing multiple opportunities for deliberative practice;
- knowing when one (teacher and student) is successful in attaining these goals;
- understanding the critical role of teaching appropriate learning strategies;
- planning and talking about teaching;
- ensuring the teacher constantly seeks feedback information as to the success of his or her teaching on the students.

The model of teaching and learning articulated throughout this chapter is based on having specific learning intentions and success criteria, as these frame the degree of challenge, the purpose, and the goals of the lesson. The common themes in what makes various strategies successful are the stipulation of planning, and in particular teachers talking with other teachers about teaching and planning, deliberate attention to learning intentions and success criteria, and a constant effort to ensure teachers are seeking feedback information as to the success of their teaching on their students. This can be enabled when teachers critically reflect on their own teaching using classroom-based evidence, and it can be maximized when teachers are in a safe and caring environment among colleagues and talking about their teaching.

Concluding comments

Teachers need to be actively engaged in, and passionate about, teaching and learning. They need to be aware of, and update their conceptions and expectations of students, and be directive, influential, and visible to students in their teaching. Teachers need to provide students with *multiple* opportunities and alternatives for developing learning strategies based on the surface and deep levels of learning leading to students building constructions of this learning. What is required are teachers who are aware of what individual students are thinking and knowing, who can construct meaning and meaningful experiences in light of this knowledge, and who have proficient knowledge and understanding of what progression means in their content to provide meaningful and appropriate feedback.

Teachers need to know the learning intentions and success criteria of their lessons,

know how well they are attaining these criteria for all students, and know *where to go next* in light of the gap between current students' knowledge and understanding and the success criteria. Teachers are successful to the degree that they can move students from single to multiple ideas then to relate and extend these ideas such that learners construct and reconstruct knowledge and ideas. It is not the knowledge or ideas, but the learner's construction of the knowledge and ideas that is critical. Increases in student learning follow a reconceptualization as well as an acquisition of information.

Enhancing learning also needs school leaders and teachers who can create school, staffroom, and classroom environments where teachers can talk about their teaching, where errors or difficulties are seen as critical learning opportunities, where discarding incorrect knowledge and understandings is welcomed, and where teachers can feel safe to learn, re-learn, and explore their own teaching knowledge and understanding. Teachers must be able to openly discuss the three key feedback questions: "Where are they going?" "How are they going?" and "Where to next?" (The "they" refers to both the teacher and to the student.)

It is also what learners do that matters. So often learners become passive recipients of teachers' lessons, but as the meta-analyses throughout this book will demonstrate, the aim is to make students active in the learning process—through actions by teachers and others until the students reach the stage where they become their own teachers, they can seek out optimal ways to learn new material and ideas, they can seek resources to help them in this learning, and when they can set appropriate and more challenging goals. Students need to be involved in determining success criteria, setting higher expectations, and being open to experiences relating to differing ways of knowing and problem solving. This then leads to their development of beliefs and reputations as a learner, and engaging in self-assessing, self-evaluating, self-monitoring, self-learning, and in learning the surface, deeper, and conceptual domains of worthwhile domains. Kember and Wong (2000) distinguished between active and passive students, and how they perceive good teaching. They found that passive learners preferred teachers who were organized, had clarity of structure, and could specify clear learning objectives, whereas active learners preferred teachers who promoted interaction in class, used a variety of teacher approaches, and displayed high levels of enthusiasm. An aim of schooling should be to maximize the number of active learners, but this requires teachers who can see learning through the eyes of their students and thence know how to engage them in learning that leads to these attributes.

As noted earlier, it is essential to have visible teaching and visible learning. This notion encapsulates directive, activating, and involved sets of actions and content, working with students so that their learning is visible such that it can be monitored, feedback provided, and information given when learning is successful. Fenstermacher and Soltis (2004) imagined the teacher as an executor, using the best learning skills and techniques available to bring about the process of learning. This is similar to the proposal by Salomon and Perkins (1989) that active learning and deep-level processing are central to success and transfer of information; or similar to the claims by Sheerens and Bosker (1997), who concluded that "it seems that highly structured learning or direct teaching, which emphasizes testing and feedback, again emerges as the most effective teaching form" (p. 219). They claimed that for transfer to occur there needs to be deep-level, connected knowledge structures—that is, knowing and understanding needs to be "conceptually deep, cohesive, and connected to other key ideas, relevant prior knowledge, multiple representations, and everyday experience" (Pugh & Bergin, 2006, p. 148). This is particularly powerful when students know

eleased under th

Document

what they know and what they do not know (the meta-cognitive awareness) and when they apply cognitive processing and meta-cognitive strategies. Motivational factors influence the success of learning due to the higher levels of engagement that thence promote these learning strategies.

The major argument is that when teaching and learning is visible, there is a greater likelihood of students reaching higher levels of achievement. It involves an accomplished teacher who knows a range of learning strategies to build on the students' surface, deep knowing and understanding, and conceptual understanding. The teacher needs to provide direction and re-direction in terms of the content being understood and thus maximize the power of feedback, and to have the skill to get out of the way when learning is progressing towards the success criteria. It also requires a commitment to seeking further challenges (for the teacher and for the student)—and herein lies a major link between challenge and feedback, two of the essential ingredients of learning. The greater the challenge, the higher the probability that one seeks and needs feedback, and the more important it is that there is a teacher to ensure that the learner is on the right path to successfully meet the challenge.

The contributions from the student

On each of about 220 days, for around 13 years, children spend five to six hours in school, nine to ten hours at home and in their communities, and about eight to nine hours asleep. When this time is added together with weekend and vacation time, students spend about 15,000 hours in school over a lifetime: or about 30 percent of their waking time is spent in the hands of those legislated to teach them. They also spend twice that amount of time (29,000 hours) at home during these school years, and they also spend 26,000 hours in the care of parents and caregivers *before* they start formal schooling (at about five to six

Student	No.	No.	No.	No.	d	SE	CLE	Rank
	metas	studies	people	effects				
Background								
Prior achievement	17	3,607	387,690	9,209	0.67	0.098	48%	14
Piagetian programs	I	51	6,000	65	1.28	_	91%	2
Self-report grades	6	209	79,433	305	1.44	0.030	102%	I
Creativity		21	45,880	447	0.35		25%	78
Attitudes and dispositions								
Personality	4	234		1,481	0.19	0.007	14%	109
Self-concept	6	324	305,859	2,113	0.43	0.010	30%	60
Motivation	6	322	110,373	979	0.48	0.047	34%	51
Concentration/	5	146	12,968	587	0.48	0.032	34%	49
persistence/engagement	:							
Reducing anxiety	4	121	83,181	1,097	0.40		28%	66
Attitude to	3	288	732,994	664	0.36	—	26%	75
mathematics/science								
Physical influences								
Pre-term birth weight	2	46	4,489	136	0.54		14%	38
Illness	2	13	—	13	0.23	—	16%t	102
Diet	I	23	—	125	0.12	0.037	8%	123
Exercise/relaxation	4	227	1,306	1,971	0.28	0.040	20%	90
Drugs	8	467	13,161	1,839	0.33	0.036	24%	81
Gender	41	2,926	5,594,832	6,051	0.12	0.034	9 %	122
Positive view of Ethnicit	y I	9	2,661	9	0.32	0.003	23%	84
Preschool experiences								
Early intervention	16	1,704	88,047	9,369	0.47	0.041	33%	52
Preschool programs	11	358	44,532	1,822	0.45	0.065	32%	55
Total	139	11,101	7,513,406	38,174	0.40	0.044	2 9 %	—

Table 4.1 Summary information from the meta-ana	lyses on the contributions from the student
-------------------------------------------------	---------------------------------------------

years of age). While the influence of schooling is probably substantial, by the time they are in their mid-teens it should also be obvious that what the child brings to the class at the start of his or her schooling, as well as on each and every day, is critical to the outcomes of education. Such out-of-school influences can come from the home, family, culture, and community. This chapter outlines some of these major attributes that the child brings to school, namely (1) background information such as prior achievement and personality dispositions, (2) attitudes and dispositions, (3) physical influences, and (4) preschool experiences.

The fundamental argument in this chapter is that students not only bring to school their prior achievement (from preschool, home, and genetics), but also a set of personal dispositions that can have a marked effect on the outcomes of schooling. While there is no doubt that schools can affect both achievement and learning dispositions the origins of both are often well in place before the child enters the school yard. For achievement, there are influences from genetics and early development, very early home and social experiences, and opportunities for learning from birth to five years (e.g., preschool and other early interventions). The key dispositional ingredients are the way the child is open to new experiences, children's emerging beliefs about the value and worth of investing in learning, and the manner in which they learn that they can build a sense of self from their engagement in the learning enterprise.

While these personality (and of course achievement) dispositions are brought by the child into the school, they also can be changed by the school—and indeed are often so changed. A major claim in this book is that schools and teachers (and researchers) may need to be more explicit that such dispositions to learning should be key performance indicators of the outcomes of schooling. Many teachers believe that if achievement is enhanced then there is a ripple effect to these dispositions. However, such a belief is not defensible, as such dispositions need planned interventions and may indeed become precursors or barriers to further learning.

As an example of the kind of dispositions that could be fostered in schools, Feist (1998) completed a meta-analysis of the differences in personality between scientists and non-scientists, creative and less creative scientists, and artists and non-artists. Creative people are more autonomous, introverted, open to new experiences, norm-doubting, self-confident, self-accepting, driven, ambitious, dominant, hostile and impulsive—a powerful cocktail. Artists have more emotional instability and are more likely to reject group norms than scientists; and creative scientists are more conscientious, conventional, and open-minded. Of these personality attributes, the attribute of openness to new experiences is critical in the success of these learners. Openness to experiences involves the willingness (and it is an active process) to experience new ideas, to think outside the box, and of not being tied to one way of thinking. It also involves a motivation to explore ideas, and to invest in the process of learning.

Openness and willingness to invest in learning is the major theme of the synthesis of meta-analyses outlined in this chapter, and this willingness to invest can be seen in many children during their first years of schooling. The continuing experience in schools can have a growing influence on this willingness to gain self-confidence and a reputation of being a learner, and such skills can be taught. Such reputation enhancement is particularly powerful during the early adolescent years and it is during these years that the "decision" to continue in education or not is often made (see Carroll, Hattie, Durkin, & Houghton, 2001). Goff and Ackerman (1992); (see also Ackerman & Goff, 1994) have explored this

notion under the heading of typical intellectual engagement and they have provided much evidence about the relationship between achievement and engagement derived from the stronger drive that some students have to invest in the development of skills and knowledge than others.

The remaining sections of this chapter refer to the four major contributions from the student: background influences, attitudes and dispositions, physical influences, and preschool effects. Each has a number of sub-categories, and for each sub-category there is a barometer console that shows the average effect size (and related information) for that particular influence; and themes are developed that are then linked together in the summary of the chapter.

Background

Prior achievement

What a child brings to the classroom each year is very much related to their achievement in previous years—brighter children tend to achieve more and not so bright children achieve less. This should not be surprising given that the correlation between ability and achievement is very high. Hattie and Hansford (1982) reported an average correlation of r = 0.51 between measures of intelligence and achievement (an effect size, d = 1.19). This high relationship accounts for what many researchers call (usually with a sense of surprise) the "Matthew effect", which is based on the biblical notion that the rich get richer and the poorer get poorer or do not gain as much. Prior achievement predicts success from preschool to the first years of schooling (Duncan *et al.*, 2007; La Paro & Pianta, 2000; Schuler, Funke, & Baron-Boldt, 1990), between high school and college or university grades (Kuncel, Hezlett, & Ones, 2001), between college and adult success (Bretz, 1989; Samson, Graue, Weinstein, & Walberg, 1984), and between grades in school and later job performance (Roth, BeVier, Switzer, & Schippmann, 1996).

Right through the education system, prior achievement is a powerful predictor. Schuler, Funke, and Baron-Boldt (1990) found that prior school grades are the best individual predictor for academic success. Fleming and Malone (1983) found that the strongest positive relationships between student characteristics and performance in science were general ability, language ability, and mathematical ability (Lapadat, 1991); and these findings were consistent across grade levels. Similarly, DeBaz (1994) found high correlations



KEY	
Standard error	0.098 (High)
Rank	14th
Number of meta-analyses	17
Number of studies	3,607
Number of effects	9,209
Number of people (8)	387,690

between science outcome measures and ability variables such as prior scholastic ability, science ability, and cognitive reasoning ability (see also Boulanger, 1981; Piburn, 1993). The overall effect size of 0.67 is among the highest effect sizes in this synthesis of metaanalyses, although the common language estimate (CLE, see Chapter 2) should remind us that, on average, prior achievement will lead to gains in achievement on 48 percent of the occasions, although there is much that is unexplained beyond prior achievement (100-48 = 52 percent that is unexplained) and so there is much that schools can influence beyond what the student brings from prior experiences and attainments. It is certainly the case that by the time the child enters school, family, preschool, or genetic factors will have already played a major role in generating subsequent differences in school-based achievement. But one of the most fascinating outcomes of this synthesis of meta-analyses is that there are measures that schools can implement that are more influential than this prior achievement effect.

Duncan et al. (2007) found that only preschool mathematics (knowledge of numbers, ordinality), and then to a lesser extent reading (vocabulary; knowing letters, words, and beginning and ending word sounds) predicted subsequent success in school. Behavior (such as externalizing and internalizing problem behaviors) and social skills were not correlated with later academic success. Although he did not conduct a meta-analysis, Feinstein (2003) reviewed the evidence on achievement before entering school using the 1970 Birth Cohort Survey, a longitudinal dataset with over 17,000 United Kingdom children. The focus was on the children's performance at 32 months, 42 months, and at five, ten and 26 years. Their measure of ability at 22 months (i.e., putting on shoes, drawing lines, pointing to facial features) was a good predictor of achievement at age 26. Children in the bottom quartile at 22 months "are significantly less likely to get any qualifications than those in the top quartile", suggesting that "before children have even entered school, very substantial signals about educational progress" are evident (Feinstein, 2003, p. 82). The effects of social class (based on parental occupation) were marked at 22 months, and if anything, the variability increased over time. This dual influence of early achievement and socioeconomic resources contribute much to what a child brings to school.

Lack of academic success

The major difference between students categorized in special education and non-special education relates (as one would expect) to achievement (Kavale & Nye, 1985; McLinden, 1988; Rush, 1992). This is not quite so clear when labels such as "at-risk" and "drop out" are used. Of the many variables that Rush (1992) investigated in his quest to see what distinguished at-risk and drop out students, the only variables where there were differences between at-risk and drop-out students compared with those not so classified were IQ, educational aspirations, and locus of control. There were differences across regions throughout the United States, pointing to the social construction of these labels rather than to some similar notion of what at-risk and drop out students are like. As these students progressed into high school other variables such as lower self-esteem and negative coping strategies become invoked to label them. This is probably because these are consequences of earlier lack of academic success, and this also then leads to attendance problems, retention, and other negative outcomes, and the use of new labels like dyslexia, ADHD, Aspergers Syndrome, and so on (the latter are real phenomena but often over used as convenient labels, Conrad, 2007; Hattie, Biggs, & Purdie, 1996).

Kavale and Nye (1985) looked at parameters of learning disabilities in achievement, linguistic, neuropsychological, and social/behavioral domains. They concluded that learning disability is a complex and multivariate phenomenon involving a number of components each making an important contribution. A comparison of learning disability grouping and normal grouping found that about three-quarters of learning disabled students could be clearly differentiated from normal students displaying deficits that would interfere with their academic ability. Sabornie, Cullinan, Osborne, and Brock (2005) used 58 studies with "high incidence disabilities" and reported large effect size differences between learning disabled children and those with mild intellectual disabilities, but a lack of difference in many school-related behaviors for these same "labeled" students. Approximately 75 percent of those labeled learning disabled and 75 percent of those labeled as having mild intellectual disabilities surpassed the average student with emotional and behavioral disabilities in achievement.

Piagetian programs

Jordan and Brownlee (1981) found that the relationship between the Piagetian stage (logical operations, concrete, formal-operational) and achievement is very high (r = 0.54, d = 1.28). This is especially the case for mathematics (d = 0.73) and it is still high but somewhat less important in reading (d = 0.40). In both subjects, seriation ability, or the proficiency to think successively (as is required to decode words on a page, count in order) was the highest correlate. Thus, knowing the ways in which they think, and how this thinking may be constrained by their stages of development may be most important to how teachers choose materials and tasks, how the concept of difficulty and challenge can be realized in different tasks, and the importance of developing successive and simultaneous thinking (Naglieri & Das, 1997; Sweller, 2008).

Self-reported grades

Another form of prior achievement is students' estimates of their own performance typically formed from past experiences in learning. Students have reasonably accurate understandings of their levels of achievement. Kuncel, Crede, and Thomas (2005) found that high school students had very accurate understandings of their achievement levels across all the subjects (r = 0.80+). This was the case for all but minority students who,





while they received lower grades than non-minority students, were more likely to be less accurate in their self-estimates or self-understanding of their achievement. Overall, however, students were very knowledgeable about their chances of success. On the one hand, this shows a remarkably high level of predictability about achievement in the classroom (and should question the necessity of so many tests when students appear to already have much of the information the tests supposedly provide), but on the other hand, these expectations of success (which are sometimes set lower than students could attain) may become a barrier for some students as they may only perform to whatever expectations they already have of their ability.

Creativity

Creativity is another prior influence on achievement, although achievement almost certainly also has a reciprocal influence on creativity (Hattie & Rogers, 1986; Kim, 2005). Murphy and Alexander (2006) evaluated the influences of knowledge, beliefs, and interests on conceptual change. The overall effects were high (d = 0.80), but greatest on domain knowledge (d = 1.31), topic beliefs (d = 0.89), concept knowledge (d = 0.69), and topic knowledge (d = 0.63). Programs with more hands-on activities had stronger effects than those relying on more passive methods such as videos or conceptual assignments. Those activities that directly addressed students' initial understandings were much more powerful than those which focused more on the presentation of accurate scientific information with less attention to students' current understandings.



KEY	
Standard error	na
Rank	78th
Number of meta-analyses	1
Number of studies	21
Number of effects	447
Number of people (1)	45,880

Attitudes and dispositions

Personality influences

There were a number of meta-analyses relating to personality influences on achievement, and to the influences on achievement of self-concept, self-estimates of ability, motivation, concentration, and engagement. The overall relationship between achievement and many of the reviewed personality variables (including anxiety, dogmatism, extraversion, locus of control, and neuroticism) is close to zero. The relationships of self-efficacy, self-concept, aspects of motivation, and persistence with achievement, however, are among the larger correlates.

O'Connor and Paunonen (2007) provide three major reasons why personality variables could have an effect on achievement. First, there are behavioral tendencies reflected in personality traits that can affect certain habits that influence academic achievement (e.g., perseverance, conscientiousness, talkativeness). Second, whereas cognitive ability reflects what an individual can do, personality traits reflect what an individual will do. Third, as students get older, personality as well as cognitive proficiency can combine to better predict subsequent performance (especially motivation-related personality variables). To assess the effects of personality on achievement, O'Connor and Paunonen related the "Big Five" factors to academic achievement. The Big Five personality factors are neuroticism (d = -0.06), extraversion d = -0.10; Boyd, 2007), openness to experience (d = 0.10), agreeableness (d = 0.12), and conscientiousness ($d \neq 0.44$, see McCrae & Costa, 1997). All correlations were small except for conscientiousness, although there was much variance for openness to experience, which suggests that there may be some circumstances when this becomes more critical to enhancing achievement. Conscientious students are thought to be more motivated to perform well academically, are typically more organized, hardworking, diligent, self-disciplined, and achievement-oriented.

Lyubomirsky, King, and Diener (2005) considered the relations between cognitive and happiness outcomes. The average effect (d = 0.54) indicated that chronically happy people and those in pleasant moods are more likely to be creative and efficient problem solvers. There is a major moderator, however, as it seems that "people in happy moods can solve complex tasks better and faster thus freeing cognitive capacity for other challenges" (Lyubomirsky, King, & Diener, 2005, p. 839). When in a "good mood we tend to make riskier judgments if nothing is at stake, but make more conservative bets when real losses are possible" (p. 839). Although achievement was not the outcome in this particular



KEY	
Standard error	0.007 (Low)
Rank	109th
Number of meta-analyses	4
Number of studies	234
Number of effects	1,481
Number of people (0)	na

meta-analysis (hence it is also not included in the tables in Appendix A), Witter, Okun, Stock, and Haring (1984) evaluated the effects of education on subjective well-being. From 556 studies, they found a mean of 0.14 for the effects of formal schooling on subjective well-being, with larger effects on life satisfaction and lower effects on happiness (see also Csikszentmihalyi, 1997; 2000; 2002).

Self-concept

I argued that self-concept related to cognitive appraisals, expressed in terms such as prescriptions, expectations, and/or descriptions that we attribute to ourselves (Hattie, 1992). Later I refined the definition in terms of a "rope" analogy. Like a rope, self-concept consists of many fibers or dimensions that intertwine and overlap rather than having any one concept of self-overpowering all (Hattie, 2005). The first premise of the rope model relates to the maxim by Wittgenstein (1958) that the strength in the rope "lies not in one fiber running throughout its length, but in the overlapping of many fibers" (Section 67). The second premise is that there are various "strands" of self-concept that serve as primary motives, that then lead to the invoking of various situation specific orientations of self (or "yarns"), such as self-efficacy, anxiety, performance, or learning orientations. In turn these situation-specific orientations lead us to choose various self-strategies ("fibers") to serve the self-motivations and thus bring meaning and predictability to our sense of selfconcept and self-esteem. Hence the rope model works as a series of interweaving threads, to form fibers and thence strands to make the rope or the sense of continuity we have of ourselves. Primary to this rope model is that self-concept relates more to how we select and interpret the information that we receive and that we present.

Teachers often makes claims about the relationship between self-concept and achievement; the common claim being that students who are high achievers have high self-concept and that it is one of their teaching roles to make students feel good about themselves such that achievement then flows. Such claims presuppose a strong relationship between perceptions of self and achievement. Hansford and Hattie (1982) looked at the relationship between various self-measures and achievement and found there was a low but positive relationship (r = 0.20); a finding that has been replicated in the United States by Holden, Moncher, Schinke, and Barker (1990), and in Europe by Muller, Gulling, and Bocci (1988). Although there is generally a stronger relationship between self-concept of ability and achievement, too often this is confounded because the self-concept of ability



KEY	
Standard error	0.010 (Low)
Rank	60th
Number of meta-analyses	6
Number of studies	324
Number of effects	2,113
Number of people (4)	305,859

measures are more self-*estimates* of ability than self-*concept* of ability (which should also include concepts of pride, worth, and confidence). The relation between self-efficacy and achievement, however, is among the strongest of self-measures (Multon, Brown, & Lent, 1991). In particular, a sense of confidence is a most powerful precursor and outcome of schooling. It is particularly powerful in the face of adversity—when things do not go right, or when errors are made. Having high levels of confidence—"can do", "want to do"—can assist in getting through many roadblocks.

Valentine, Du Bois, and Cooper (2004) conducted a meta-analysis of three causal models that have been proposed to account for the relationship between self-concept and achievement: (a) the skill development model, which suggests that student achievement causes self-concept; (b) the self-enhancement model, which suggests that student self-concept causes achievement; and (c) the reciprocal effects model, which suggests that student student achievement and self-concept affect each other in a reciprocal fashion. They found that there was more support for the reciprocal effects model of causal relations between self-concept and achievement than for any of the other models examined. Valentine, Du Bois, and Cooper (2004) concluded that "these results lend further support to social cognitive theory, specifically that affective, cognitive, and environmental variables interact in a reciprocal fashion to determine human behavior" (p. 28). This should hardly be surprising given the low covariance between self-concept and achievement.

Further such investigations of causality are unlikely to resolve the directionality. Instead, it is more likely that there are stronger relationships between certain self-strategies and achievement. Achievement is more likely to be increased when students invoke learning rather than performance strategies, accept rather than discount feedback, benchmark to difficult rather than to easy goals, compare themselves to subject criteria rather than to other students, possess high rather than low efficacy to learning, and effect self-regulation and personal control rather than learned helplessness in the academic situation. The will-ingness to invest in learning, to gain a reputation as a learner, and to show openness to experiences are the key dispositional factors that relate to achievement. Maybe it is the choice of these strategies, not the level of self-concept, that is the precursor to achievement gains, and it is likely that success in achievement also reinforces the choice of these self-strategies. Maybe it is, therefore, not surprising that teachers have more difficulty changing the levels of achievement of those with non-supportive self-strategies; they may have more success if they addressed these strategies *before* attempting to enhance achievement directly.

Motivation

In the 1960s, the British philosopher Richard Peters (1960) challenged the value of the concept of motivation. He argued that the concept of motivation implied a push or pull notion, whereas children make decisions to do this rather than that all the time. Children are moving anyway so discussion about pushing or pulling implies a false assumption of a static being. Indeed, a major mission in education is to ask "Why math rather than billiards?", "Why spend effort on homework and not baseball?", "Why learn more when I know enough to pass?" The minimax principle—minimal effort for maximum gain—can be most strategic but hardly enhancing. Schools, however, tend to always ask for more and the overbearing request for more is often resisted by students who hear such demands from every teacher. A key aspect in the discussion about motivation needs to relate to the



purposes and goals, the learning intentions and challenges, and the personal strivings of students, as much as it needs to relate to the intrinsic properties of the task and who makes the demands.

There is much value anticipating when student motivation is at its highest. Dörnyei (2001) noted that motivation is highest when students are competent, have sufficient autonomy, set worthwhile goals, get feedback, and are affirmed by others. He also challenged educators to seriously consider student demotivation caused by, for example, public humiliation, devastating test results, or conflicts with teachers or peers. For many, demotivation has more impact than motivation. Such demotivation can directly affect commitment to the goals of learning, turn off the wish for and power of feedback, and decrease involvement. It can take less effort by a teacher to demotivate students compared to the often greater effort required to motivate them—to turn students on to learning.

Having a sense of control over one's learning can be important. Ross (1988) reviewed the evidence on the degree to which students learnt how to control their own learning (when completing science studies) and this was highly related to outcomes. This sense of control over one's learning, or a "person's beliefs about control over life events" (Findley & Cooper, 1983, p. 419) has been often studied. Students who take on personal responsibility for life events such as learning can be labeled *internals*, whereas those who consider learning are out of their hands are externals. The typical finding is that more internal beliefs are associated with academic achievement. The influence is greater for males than females, and more so for adolescents than for children or adults (Findley & Cooper, 1983; Kalechstein & Nowicki, 1997), although some have reported no such differences (Sohn, 1982). In their meta-analysis, Frieze, Whitely, Hanusa, and McHugh (1982) found that males make stronger ability attributions regardless of the outcome, whereas females have a slight tendency to attribute failure to luck. The notion that increasing achievement is a function of our efforts and interest is critical to success-there is no point, for example, in investing in study or preparation if we do not believe that our efforts can make a difference. Certainly interest plays a part in choosing subjects and choosing to commit to expending effort, and, as Schiefele, Krapp, and Winteler (1992) discovered, interest is also related to achievement (d = 0.62). The effects for interest and achievement were greater for females (d = 0.70) than males (d = 0.50), for the natural (d = 0.68) compared to the social sciences (d = 0.48), but similar across grade levels. Twenge, Zhang, and Im (2004) found that there has been an increase over the past two generations in students claiming that learning is more external than internal. They argued that students were becoming more cynical, and

using more ineffective stress management strategies, and it could be suggested that the increased emphasis on external accountability testing models has not helped.

Concentration, persistence, and engagement

Engendering a positive attitude to school work may be both a precursor to greater engagement, and a worthwhile outcome in itself. It seems achievement plus effort plus engagement are keys to success in school. We should not make the mistake, however, of thinking that because students look engaged and appear to be putting in effort they are necessarily achieving; this is one of the myths that is held in too many classrooms—busy work alone does not make the difference. The discussion throughout this book about clear learning intentions, transparent success criteria, and making learning visible to the student are the key elements of engaging students.

Engagement in Kumar's (1991) meta-analysis was defined as the effective time within allocated science class that a student actively participated in learning—such as experimenting, attending, participating in discussion, questioning, answering, and taking notes. The overall effects were very high indeed (d = 1.09). Similarly, there were high relations between engagement and degree of concentration on tasks (Datta & Narayanan, 1989). Feltz and Landers (1983) showed that one way of enhancing concentration is to mentally visualize the processes and strategies involved in a task: students who mentally visualized various motor tasks were more effective compared to those that did not (d = 0.48).

These effects of engagement and concentration seem to be similar across ethnic groups. Cooper and Dorr (1995) found that there were no differences between African American and white students in their need for achievement, personal expectations, feelings of hopelessness, denial of the importance of individual effort, or lack of persistence.

Reducing anxiety

Spielberger (1972, p. 1) described anxiety as the outcome of a "chain reaction consisting of a stressor, a perception of threat, a state reaction, cognitive reappraisal, and coping". The meta-analysis research in education often focuses on two prominent forms of anxiety: test anxiety and mathematics anxiety. The subject of mathematics in particular promotes expressions of anxiety that take such forms as tension and dislike (attitudinal features); worry, helplessness, and mental disorganization (cognitive features); and fear (emotional



0.032 (Low)
49th
5
146
587
12,968



feature). The consequences of anxiety include avoidance of courses and an inability to achieve in the subject (Ma, 1999). The four meta-analyses discussed here were concerned about the effects of anxiety on achievement, although the effects have been reversed to indicate the gain in achievement that can occur if this anxiety is reduced.

Seipp (1991), for example, reported an effect size of -0.43 (r = -0.21) between anxiety and performance and noted it was similar for males and females. Worry (d = -0.44) was more negatively related to emotionality (d = -0.30), and test anxiety (d = 0.46) had greater debilitating effects than general anxiety (d = -0.32). Hembree (1988) established that test anxiety was significantly related to achievement for students from grade 3 and above. Relationships were stronger for worry than emotionality. Those students with high or low (as opposed to middle) self-concept tended to be more test-anxious, and there were direct relationships to students' fears of negative evaluation, defensiveness, and dislike of tests. Some specific attributes of tests that can invoke higher levels of anxiety include the use of "none of the above" as a multiple choice option, distorted pictures with word problems, and the presence of extraneous information in word problems (Hembree, 1987; Ma, 1999). Teachers need to consider methods to reduce anxiety, as it can be an important barrier to learning.

Attitude to school subjects

Attitude to school involves many dimensions, such as positive or negative feelings, the tendency to engage in or avoid school activities, a belief that one is good at schoolwork



or not, and a belief that school is useful or not (Aitken, 1969; Ma & Kishor, 1997; Neale, 1969). The effects of attitude to mathematics on mathematics achievement are as substantial as the more generalized personality variables cited above, and potentially more amenable to teacher influences (Ma & Kishor, 1997). There are no major differences relating to sex or grade level (Ma & Kishor, 1997), although girls show slightly more anxiety to mathematics (d = -0.18), and less self-confidence in mathematics (d = -0.12, Etsey & Snetzler, 1998).

Although developing attitudes towards school and subjects is a desirable outcome of schooling, it clearly is also a correlate of achievement and it is suggested that by enhancing attitudes there could be reciprocal effects on achievement.

Physical attributes

There are a number of background factors that can affect children before they come to school—some are out of the control of the child, for example birth weight or illness, and some are more related to nutrition, exercise, and the use of drugs. Two of the physical attributes are among the most discussed moderators to performance: gender and ethnicity.

Birth weight

Bhutta, Cleves, Casey, Cradock, and Anand (2002) showed that birth weight of pre-term born children was associated with lower cognitive scores at school age compared with the birth weight of full-term-born children. They found that the cognitive scores of preterm cases and term-born cases were directly proportional to their birth weight. The typical age of measurement of cognitive skills was eight to ten years. They argued that this decrease was not surprising given the developmental vulnerability of the immature brain, and factors such as severity of illness in pre-term neonates, their physiological instability and exposure to early adverse experiences. Of course, there are limitations (as noted in the article) of compounding effects such as socioeconomic, nutritional, and other family factors that could moderate these conclusions.

Corbett and Drewett (2004) investigated those children who failed to thrive in the early days and months, and while the effect size was not as substantial as with the pre-term babies, failure to thrive in infancy was associated with adverse outcomes at a later age. It is



KEY	
Standard error	na
Rank	38th
Number of meta-analyses	2
Number of studies	46
Number of effects	136
Number of people (2)	4,489



extremely likely, however, that there are many successful interventions to ameliorate these early influences.

Illness

The effects of chronic illness on achievement are negative but small (Sharpe & Rossiter, 2002, d = -0.20). When parents were asked to rate the chronicity of the illness, they were much more negative than when children rated the impact of the illness on their achievement, possibly because they are more protective and children may not perceive as many negative effects as being directly related to their illnesses. The negative effects pertained not only to cognitive outcomes (d = -0.20) but also to peer activities (d = -0.29), psychological functioning (d = -0.22), but less so to self-concept (d = -0.06). There was not much difference between the chronic illnesses: cancer d = -0.28, diabetes d = -0.23, anemia d = -0.26, bowel disorders d = -0.32, and spina bifida d = -0.26, but cardiac illness was d = 0.20. These differences could also reflect absences from school.

Diet interventions

There have been many arguments that the eating of certain foods or the presence of food additives affects students' achievement. Kavale and Forness (1983), in a meta-analysis looking at hyperactivity and diet treatment, found that the Feingold and Feingold (1979) claim that reduction in artificial food additives (colors and flavors) is not an effective



KEY	
Standard error	0.037 (Low)
Rank	123rd
Number of meta-analyses	1
Number of studies	23
Number of effects	125
Number of people (0)	na

intervention for hyperactivity. The negligible treatment effects from such interventions were only just greater than those expected by chance. Similarly Kavale and Dobbins (1993) found that diet interventions have a limited effect in terms of student behavior despite favorable public responses.

Exercise and relaxation

The relation between physical fitness and exercise on cognitive functioning is small but positive (Etnier, Nowell, Landers, & Sibley, 2006; Etnier *et al.*, 1997). The length or intensity of fitness programs did not have a differential effect on cognitive functioning, but relaxation techniques in general, and progressive relaxation techniques in particular, had a small positive effect on cognitive academic variables among elementary school and college level students (Moon, Render, & Pendley, 1985). Dishman and Buckworth (1996) found that intervention programs based on forms of behavior modification were associated with larger effects of physical exercise and achievement. Interventions in community settings and interventions delivered to groups reported larger effects, contrasted with those in schools and other settings or with delivery to individuals and the family. Effects were larger when the physical activity was not supervised compared with a supervised physical activity program. Effects were unrelated to the number of weeks of the intervention or the follow-up period.

Strong *et al.* (2005) identified 850 articles on the effects of physical activity on health and behavior outcomes among school age students. Most interventions used supervised programs of moderate to vigorous physical activity of 30 to 45 minutes for three to five days a week. They found that the addition of physical education to the curriculum resulted in small positive gains in academic performance, and, as important, allocating time away from other subjects to physical education did not detract from achievement in other subjects. The effects came mainly from small positive effects on concentration and memory, and enhanced classroom behavior.

Drug interventions

So often drugs are seen as the answer to reducing behavior problems and for enhancing attention, and there are claims they can increase achievement. The evidence is more equivocal. Purdie, Hattie, and Carroll (2002) investigated the effects of various drugs on



KEY	
Standard error	0.040 (Low)
Rank	90th
Number of meta-analyses	4
Number of studies	227
Number of effects	1,971
Number of people (1)	1,306



students with attention deficit/hyperactivity disorders (ADHD/ADD) and while there was evidence that the various drugs (both stimulants and depressives) seemed to reduce behavior problems at least when rated by teachers and parents, although not by the students or independent observers (see also Crenshaw, 1997), there were limited effects on achievement. Crenshaw (1997) found an effect of d = 0.52 for classroom tests versus d = 0.25 for standardized achievement tests. In her meta-analysis, the effects on behavior (d = 0.68) were much higher than on achievement (d = 0.90) and increased attention span (d = 0.84) compared to academic performance (d = 0.47, see also Silva, Munoz, & Alpert, 1996; Thurber & Walker, 1983).

DuPaul and Eckert (1997), however, found that school-based programs had a greater effect on the achievement of ADHD students compared to drug treatments. They investigated in-school treatment programs for ADHD students and their effect on behavior was d = 0.78and achievement d = 0.58. Contingency management (d = 0.94) and academic interventions (d = 0.69) were more effective than cognitive-behavioral procedures (d = 0.19) in improving classroom behavior. The latter, they argued, is most effective in enhancing achievement effects.

Kavale (1982) found similar positive effects for stimulants but more so for lower level tasks (memory and copying, d = 0.41) than for higher order tasks (reading d = 0.32 and mathematics 0.09). In Purdie *et al.* (2002), it is worth noting that the effects of stimulants (d = 0.35) was not that different from school-based psychological and educational interventions (d = 0.39), social skills training (d = 0.31), cognitive and self-regulation programs (d = 0.58), and parent training (d = 0.31).

It seems that there is a syllogism in play here: drugs reduce behavior problems; when problem behaviors are reduced students are more likely to be attentive; when a student is attentive they may learn. Too often, the conditional "mays" are ignored and the straight causal connection made. While not denying that children with ADHD and other medically derived conditions exist, there does need to be concern about the pathologizing of barriers to learning. The concern for schools is to find teaching and learning processes such that whatever the etiology of non-learning, the aim is to allow students to learn. There are many successful strategies for teaching students to attend, to develop social skills, and to participate in learning, and labels should not be the excuse for why schools are not successful. All students arrive at school unique, and whatever their differences our aim in schooling is to provide optimal conditions for success in
41

learning and not use labels to justify why these students cannot be as successful as any others (Conrad, 2007).

Gender

There is a received wisdom for many in education that there are marked differences in the achievements of males and females. Much of our popular messages are that "men are from Mars, women are from Venus" (Gray, 1993), or "boys are for mathematics and girls are for language"-there are many claims that emphasize the differences between the sexes (single-sex classes or schools, different programs for girls and boys, and so on). The predominant message from the synthesis of meta-analyses, however, is support for lanet Hyde's (2005) argument about gender similarities. This argument proposes that males and females are similar on most, but not all, psychological variables. They are more alike than they are different. The evidence for this claim is overwhelming.

Hyde (2005) collected 124 meta-analyses across many psychological dimensions, and although achievement is the major interest in this book, the message in most other areas is also one of similarity between males and females. Differences in communication (interruptions, talkativeness, self-disclosure, facial expression processing) were small (d = -0.15in favor of females), differences in social and personality variables (aggression, negotiation, helping, leadership, extraversion) were small (d = 0.18 in favor of males), and differences in well being were small (d = -0.06 in favor of females). Larger differences were noted as exceptions to this similarity message. Males outperformed females in motor performance and physical aggression, and females outperformed males in agreeableness. Hyde also considered nine meta-analyses concerning achievement outcomes (all included in the current synthesis) and reported an effect size of d = -0.06 (in favor of females). In the current synthesis, the average across 39 meta-analyses and 2745 effect sizes is d = 0.12(in favor of males). The overall differences are small and the gender similarity hypothesis advocated by Hyde is much supported.

The only question, therefore, is why we are so constantly immersed in debates about gender differences in achievement-they are just not there. The current synthesis shows that where differences are reported, they are minor indeed. For example, while sex differences are virtually zero in verbal ability (Hyde, 1981; Hyde & Linn, 1988) there are very small differences in mathematics (Freeman, 1984; Friedman, 1989; Frost, Hyde, & Fennema, 1994; Hines, 1989; Hyde, Fennema, & Lamon, 1990; Hyde, Fennema, Ryan, Frost, &



Hopp, 1990; Linn & Hyde, 1989), and very small differences in science (DeBaz, 1994; Kahl, Fleming, & Malone, 1982; Steinkamp & Maehr, 1984). Girls' motivational orientation was more positive in biological sciences and chemistry, whereas boys outscored girls in physical sciences (Becker & Chang, 1986; Haladyna & Shaughnessy, 1982; Steinkamp & Maehr, 1984; Weinburgh, 1995). Whitley (1997) showed that males see computers as more suited to themselves than do females. Males also see themselves as more competent in computerrelated tasks and have higher self-efficacy. Most differences were small.

Cohn (1991), examining sex differences in personality development, found that adolescent girls achieve developmental milestones, including ego development, earlier than boys, but that these differences declined with age. These differences are relatively stable during early and middle adolescence, moderately larger among junior and senior high school students, declining significantly among college-age adults, and disappearing among older men and women. In a meta-analysis of gender differences in temperament, Else-Quest, Hyde, Goldsmith, & Hulle (2006) found that girls have slightly higher scores on attention (d = 0.23) and persistence (d = 0.08), and very large differences on effortful control (d = 1.01) and inhibitory control (d = 0.41). Thus girls "display a stronger ability to manage and regulate their attention and inhibit their impulses" (p. 61).

Lytton and Romney (1991) used meta-analysis to investigate whether parents have systematic differences in their patterns of rearing boys and girls. Effect sizes in most socialization areas are small. In North American studies, the only area displaying a significant effect is the encouragement of sex-typed activities such as play and household activities by parents; physical punishment is applied significantly more to boys, and fathers differentiate more than mothers between boys and girls.

A meta-analytic review of gender differences in group performance (Wood, 1987) showed that while all-male groups performed better than all-female ones, the differences once again were small. Female group members' interaction facilitates performance at tasks requiring positive social activities including friendliness and agreement with others, whereas males' styles of interaction facilitate performance on tasks requiring task-oriented behavior such as giving opinions and suggestions.

A related attribute, leadership style, has been studied more in adults (e.g., principals) than school-age students. Pantili, Williams, and Fortune (1991) looked at the effectiveness of assessment by the National Association of Secondary Schools Principals in evaluating desirable criteria for principalship. Criteria such as sensitivity, range of interests, and personal motivation have almost no effect on job performance. Indeed, neither gender nor ethnicity has any significant effect on the assessment centre scores of principals, on any dimension. Eagly, Karau, and Johnson (1992) reviewed 50 studies that compared gender and leadership style in principals of public schools. The most substantial gender difference is the tendency for female principals to lead in a more democratic, less autocratic style and tend to lead in a more task-oriented way than male principals. Men adopt a less collaborative style and are relatively more dominating and directive than women.

Overall, the differences between males and females should not be of major concern to educators. There is more variance within groups of boys and within groups of girls than there are differences between boys and girls. Hyde (2005) noted, for example, that there was no evidence for Gilligan's (1982) claims that women speak in a different moral voice of caring and men in terms of justice, and that there is therefore no reason to believe

men or male teachers are not nurturing and girls motivated by a sense of justice. Similarly mathematically talented girls and reading talented boys can be overlooked given the adults' (parents and teachers) beliefs and expectations about sex differences in these areas.

Ethnicity

It was possible to find only one meta-analysis specifically related to ethnic differences and achievement (although the research on desegregation had a focus on ethnicity; see below). The focus of the study by Allen, Bradford, Grimes, Cooper, and Howard (1999) was on racial group orientation and social skills, although they did include achievement as one of their outcomes. Racial group orientation is the degree to which a student has a positive view of his or her own ethnicity. Such students demonstrate enhanced academic success (d = 0.32), an increase in positive developmental adjustment (d = 0.40), a decline in participation in delinquency (d = -0.23), and an improvement in sociability (d = 0.30). It certainly seems that maintaining a positive image of our cultural background is most worthwhile.

There has been no meta-analysis, however, exploring differences in ethnicity and achievement. It seems remarkable that one of the more important moderators of influences on achievement has not been the subject to meta-analytic exploration. Within some meta-analyses (but not as many as I would have expected) ethnicity is used as a moderator. There are no differences in effect sizes for the amount of formal schooling (Willig, 1985, African American d = 0.18, white d = 0.16); the presence of fathers or not (Schneider, 1992, African American d = 0.25, white d = 0.25; Albanese & Mitchell, 1993, African American d = 0.75, white d = 0.98), and no differences in need for achievement (Cooper & Dorr, 1995, white d = 0.02 greater than African Americans for studies post-1970). The only difference related to small group learning (Evans & Dion, 1991, white d = 0.48, African American and Hispanic d = 0.97).

There is no reason, from this limited number of effect sizes, to believe African American and white students are differentially affected by what works best or by the underlying features of the model outlined in Chapter 3. What seems more important is that students have a positive view of their own racial group, and that educators do not engage in the language of deficit theorizing. Accepting that students come to school with different cultural heritages and that they can be allowed and encouraged to have



KEY	
Standard error	0.003 (Low)
Rank	84th
Number of meta-analyses	1
Number of studies	9
Number of effects	9
Number of people (1)	2,661

a positive image of their own racial or cultural heritage is an acknowledgment of the importance of culture, and can show the students that they are accepted and welcomed into the learning environment (see Bishop, 2003). Further, so much discussion is about the tail or gaps between white students and students of color—but such language is misleading as there are many gaps in achievement for students of all ethnicities, both above and below the mean of achievement. So often only the gaps below the mean are considered, and worse, generalized as if all students are near the bottom of the distribution (see Hattie, 2008).

Preschool influences

Early interventions

One of the claimed keys to success at school is the amount and nature of preschool experience that a child has before starting school. The overall effect of early intervention (any intervention with preschool age students) is d = 0.50, and for preschool programs (a specific program such as kindergarten) is d = 0.52. The overall finding is that early intervention programs are more effective if they are structured, intense, include about 15 or more children, and the children are in the program for up to 13 hours a week. This effect accrues similarly for regular students as well as for at-risk, disabled, and special education students. The effects, however, reduce over time and thus there is a need for systematic, sustained, and constant attention to enhancing learning if these early gains are to be maximized.

The benefits of early intervention are evident over a variety of outcome variables (including IQ, motor skills, language, and academic achievement) and across a wide variety of children, conditions, and types of program. The best early predictors of achievement in these meta-analyses are attention distractibility, internalizing behavior problems, language variables, and tests of general cognitive functioning (Horn & Packard, 1985). The more effective programs are more highly structured, and run by well trained staff (Innocenti & White, 1993). There is little support for the widely held belief that involvement of parents leads to more effective early intervention (Casto & Mastropieri, 1986; Casto & White, 1985; K. R. White, 1986), although there is support for the claim, however, that those most in need (disadvantaged students, for example, students from lower socioeconomic areas, or minority students) gained the most (Collins, 1984; Harrell, 1983).



KEY	
Standard error	0.041 (Low)
Rank	52nd
Number of meta-analyses	16
Number of studies	1,704
Number of effects	9,369
Number of people (5)	88,047

Preschool programs

Goldring and Presbrey (1986) completed a meta-analysis of preschool intervention programs for disadvantaged children and found that preschool intervention programs do have positive effects regardless of diversities in sites, length of intervention and curriculum models. Children who took part in preschool intervention programs were still showing, in elementary school, a gain of about half a standard deviation more than counterparts who had not taken part in such programs. By high school the gain was negligible. Jones (2002) found that all-day kindergarten had high effects (d = 0.56) on achievement in early school—with the greatest effects on reading and language rather than on mathematics (d = 0.60 compared with d = 0.40). La Paro and Pianta (2000) also reported similar effects of d = 0.43 between academic scores in preschool to kindergarten and d = 0.48 between kindergarten to first and second grade for academic outcomes (and d = 0.32 and d = 0.29 respectively for social outcomes).

The type of preschool seems to be an important moderator. Fusaro (1997) found that children attending full-day kindergarten showed significantly greater achievement than those attending half-day. Applegate (1986) reported negative effects for day care compared to home care children on attachment directed towards parent, but they were less frustrated, cried less often, were less tense and showed less attachment towards a non-parent figure, had a greater increase in exploratory behavior, and were less often reprimanded. Day care children showed greater gains on cognitive areas (d = 0.43), emotional (d = 0.56), and social/behavioral (d = 0.04) compared to home care children.

There is little evidence that earlier is better in starting intervention programs, and any effects decline quickly over time (Casto & Mastropieri, 1986; K. R. White, 1986). For example, for disadvantaged populations, the immediate benefits decline rapidly and largely disappear after 60 months (Casto & Mastropieri, 1986; Casto & White, 1984; Kim, Innocenti, & Kim, 1996; White & Casto, 1985; White, 1986). Gilliam and Zigler (2000) synthesized the effects of preschool across 13 American states, and claimed there were sizable effects (d = 0.2-0.3) on achievement by the end of preschool, although these effects were not evident by the end of first grade.

Nelson, Westhues, and MacLeod (2003) reported that the effects of these preschool programs were greater if students were in them for at least a year, and were particularly higher for minority students. In mathematics, performance on standardized mathematics (d = 0.25) and reading (d = 0.20) was higher for participating than for non-participating children. By the upper grades (grades 7–11), a slightly higher percentage of underachieving



KEY	
Standard error	0.065 (Medium)
Rank	55th
Number of meta-analys	ses 11
Number of studies	358
Number of effects	1,822
Number of people (4)	44,532

eled

students who had participated in preschool intervention programs did not need special education and were not held back a grade (Goldring & Presbrey, 1986).

Concluding comments

A common theme throughout this chapter is the powerful influences that the child brings into the school—via the effects of their prior achievements, their personality dispositions, and their preschool experiences. Equally noteworthy is the low to non-substantial effects of gender, diet, and exercise. While the very earliest influences of prior achievement (whatever is gained via genetics, early parenting, or preschool intervention) may be least influenced by school education, there are many opportunities throughout school to influence some of the key attributes, such as the willingness to engage in learning, the degree of reputation enhancement that a child can gain from being engaged in learning, the attributions of success to factors such as effort rather than ability, and the raising of positive attitudes towards learning.

Some of the most fascinating and important influences—openness to experience, willingness to invest in learning, and intellectual engagement—can be fostered in preschools, and then developed to a particularly high level in our schools. This can be done by ensuring that tasks are appropriately challenging to students, and that success is attributed to their investment in the tasks. This can then lead to a sense of reputation enhancement, whereby students derive a sense of self and reputation among peers that they are learners. Therein lies success. Such personal dispositions can have a marked effect on the outcomes of schooling. It is also worth noting that while many personality variables increase after students leave compulsory schooling (such as social dominance, agreeableness, conscientiousness, and emotional stability), this is not the case with openness to experience—which is one of the more powerful influences on achievement throughout schooling (Roberts, Walton, & Viechtbauer, 2006).

Many preschool programs can have an effect on these dispositions, as does the early development of successive processing skills, which can give children a head start in the achievement stakes. These successive skills, such as seriation, decoding sequences, and seeing from the parts to the whole appear to be given a major boost by participating in early intervention programs (Naglieri & Das, 2001; Luria, 1976). Students who have begun to master these successive processing skills before they enter formal schooling have an advantage when they begin the more formal learning.

The contributions from the home

The home can be a nurturing place for the achievement of students, or it can be a place of low expectations and lack of encouragement in learning. Most parents, however, begin with positive aspirations for their children: certainly children are born into a set of expectations and these expectations can be critical to the success of children when they go to school. A major concern is that some parents know how to speak the language of schooling and thus provide an advantage for their children during the school years, and others do not know this language, which can be a major barrier to the home making a contribution to achievement. This chapter investigates the influences of the family resources, the family structure and environment, television, parental involvement, and home visiting. A theme developed in this chapter is that parents can have a major effect in terms of the encouragement and expectations that they transmit to their children. Many parents, however, struggle to comprehend the language of learning and thus are disadvantaged in the methods they use to encourage their children to attain their expectations.

Socioeconomic status

Socioeconomic status (SES) relates to an individual's (or family's, or household's) relative position in the social hierarchy and directly relates to the resources in the home. Such resources refer to parental income, parental education, and parental occupation as three main indicators of SES. The overall effect from the four meta-analyses based on 499 studies (957 effects) is d = 0.57, which is thus a notable influence on the student's achievement.

Home	No. metas	No. studies	No. people	No. effects	d	SE	CLE	Rank
Socioeconomic status	4	499	176,915	957	0.57	0.016	40%	32
Welfare policies	1	8	_	8	-0.12	0.030	-8%	135
Family structure	13	845	10,147,912	1,733	0.17	0.032	12%	113
Home environment	2	35	5,83 I	109	0.57		40%	31
Television	3	37	1,022,000	540	-0.18		-12%	137
Parental involvement	11	716	320,000	1,783	0.51	0.178	36%	45
Home visiting	2	71	_	52	0.29	_	20%	89
Total	36	2,211	11,672,658	5,182	0.31	0.058	22%	—

Table 5.1 Summary information from the meta-analyses on the contributions from the home

But it is important to consider the influences of these various sub-components of SES before discussing its effects as if it were a unidimensional notion.

In the meta-analysis of 58 studies by Sirin (2005), the effect size between achievement and parental education was d = 0.60, parental occupation was d = 0.56, and parental income was d = 0.58: very similar indeed. Further, there was an effect size of d = 0.50with neighborhood resources, and d = 0.66 with free or reduced cost lunches (a common measure of SES in the US). There was little variability in the relation between SES and various types of achievement (verbal d = 0.64; mathematics d = 0.70; science d = 0.54). Sirin made much of a slight increase from pre-school through middle school, but the effects are not that different: pre-school d = 0.38, elementary d = 0.54, middle d = 0.62, and high school d = 0.52. The effects were lower for students in rural schools (d = 0.34, where there is likely less variability of SES within a school) than in suburban (d = 0.56) and urban (d = 0.48) schools. Overall, there were not many differences across these effects based on the key components of SES, so the question arises as to how these SES effects influence student achievement.

It is likely that the effects from socioeconomic resources are more influential during the pre-school and early years of schooling. For example, Hart and Risley (1995) showed that when students from lower SES groups start school, they have, on average, spoken about 2.5 million words, whereas those from higher groups have spoken 4.5 million words: this demonstrates a remarkable difference in what students bring to school. The lack of resources, the lower levels of involvement in teaching and schooling, the lesser facilities to realize higher expectations and encouragement, and the lack of knowledge about the language of learning may mean that students from lower SES groups start the schooling process behind others.

We need to be careful, however, about the unit of analysis used in these studies: is it the socioeconomic status of the school or of the student? White's (1982) meta-analysis on the relationship between SES and academic achievement noted the importance of distinguishing between effects based on aggregated units (such as SES of the school) versus effects based on the individual level (such as the SES of the student). The aggregate effect was d = 0.73 at the school level, whereas the effect was d = 0.55 at the individual student level. Further, Sirin noted that the effect was much lower when the data about SES were provided by the students (d = 0.38)—who probably saw less inequity in the difference due to home resources—than when provided by the parents (d = 0.76).



KEY	
Standard error	0.016 (Low)
Rank	32nd
Number of meta-analyses	4
Number of studies	499
Number of effects	957
Number of people (2)	176,915

SES is more important at the school than at the individual level, and for the parents more than for the students. This raises the question of the notion of adequacy of funding at the school level—that is, the sufficiency of resources for optimal academic achievement rather than equity, which usually means smoothing the differential resources at the student or family level but not acknowledging the increased level of problems and issues faced by schools teaching students from poorer backgrounds. A criticism of much of the school effectiveness literature is that the cultures and sub-cultures within schools are often left out (Slee, 1998). Certainly the culture and politics of schools have a major role in explaining why a school is or is not effective. A major premise of this book is that the visibility of teaching and learning is indeed a within-school phenomena, can be encouraged or discouraged by the culture and politics within schools, and probably can only be maximized as a function of within-school cultures and politics.

One of the ways this influence is manifested is that schooling introduces a language and set of cultural norms with which many parents, particularly those from lower SES families, are not familiar. In a five-year evaluation of five of the lowest SES schools in New Zealand, we found major consequences when teaching parents the language of schooling (Clinton, Hattie, & Dixon, 2007). This evaluation of what was known as the Flaxmere Project involved a series of innovations related to improving home-school relations within and between these five schools, including giving families computers and employing former teachers as "home-school liaison persons". The home-school liaison persons allowed the parents to learn the language of schooling-that is, the parents learned the language about the nature of learning in today's classrooms, learned how to assist their children to attend and engage in learning, and learned how to speak with teachers and school personnel. Involving parents with the schools via the Flaxmere Project led to enhanced engagement by students in their schooling experiences, improvements in reading achievement, greater skills and better jobs for the parents, greater awareness of the language of schooling, and higher expectations, high satisfaction, and high endorsement of the local schools and the Flaxmere community (the effect sizes ranged from d = 0.30-0.60 and occasionally were much higher across many outcomes). The greatest effects were an increased knowledge of the language of schooling and learning by the parents.

Either there can be efforts to reduce the barriers between school and home or the effects of the home on student learning can be compromised as the child is then asked to work in two worlds—the world and language of home, and the world and language of school. For many children this is asking too much. It is also difficult for children in these two worlds to build a reputation as a learner, learn how to seek help in learning, and have a high level of openness to experiences of learning.

Welfare policies

Gennetian, Duncan, Knox, Clark-Kauffman, and London (2004) found in their meta-analysis close to zero effects from students in families who received welfare compared to those not receiving welfare. While they make much of an effect size of d = -0.10, by claiming that the effects on adolescents were "significantly worse", it is difficult to imagine the visible effects of findings such as about four percent fewer mothers in the welfare program group reporting that their child performed above average, and only about two percent more of this group of mothers indicating that their child repeated a grade. There are certainly



many other effects of welfare programs for these families that are beneficial, but it seems that there are other more powerful effects on achievement than the welfare status of the family.

Family structure

There are many types of families, and the effects of these different types could be classified as small compared to many other influences.

Single and two-parent families

About 70–80 percent of families have two parents in most Western countries, about 10–20 percent of families are single-parent, and about 2–10 percent are other than these structures. Pong, Dronkers, and Hampden-Thompson (2003) found that single parenthood is associated with lower mathematics and science achievement (although the effects are quite small). They also noted that countries with more generous welfare policies, like Austria, showed the smallest gaps. The greatest gaps were in countries such as the United States and New Zealand, who, they claimed, lagged behind other industrialized countries in providing financial assistance, universal child benefits, tax benefits and maternity leave benefits to single and poorer families. They concluded that "to some extent the investment in national family policies explains why Australia ranks at the top but the United States and New Zealand rank last in the academic resilience of children from single-parent homes" (p. 695).



KEY	
Standard error	0.032 (Low)
Rank	113th
Number of meta-analyses	13
Number of studies	845
Number of effects	1,733
Number of people (3)	10,147,912

Resident and non-resident fathers

The three meta-analyses on this topic all found small effects on achievement relating to whether a father was present or not in the family. Amato and Gilbreth (1999) found small effects relating to the fathers paying or not paying child support (d = -0.13), contact between fathers and children (d = 0.11), feeling close (d = 0.06), and authoritative parenting (d = 0.17). Salzman (1987) found a d = 0.26 achievement effect of father-presence compared to father-absence. The effects were slightly higher for achievement (d = 0.30) compared to aptitude tests (d = 0.20), for elementary and junior high students than for pre-school, but there were no differences for males and females or across socio-economic groups.

Divorce

Compared to children with continuously married parents, children with divorced parents scored lower (but not by much) on measures of academic achievement, psychological adjustment, self-concept and social relations. Amato and Keith (1991) used 92 studies that compared children living in divorced single-parent families with children living in continuously intact families. The overall effect size was d = 0.16 lower on school achievement for the children in the former group although this difference was lower for the more recent studies (d = -0.12 for more recent studies, compared to d = -0.23 for studies 30 or more years ago). Other effects were d = -0.23 for conduct, d = -0.08 for psychological adjustment, d = -0.09 for self-concept, d = -0.12 for social adjustment: all small effects. Teachers saw no differences between these two sets of children (d = -0.04), and the effects were greater for girls than boys (d = -0.30). Amato and Keith also found similar achievement effects (d = -0.22) for children who experienced the death of a parent.

Jeynes (2006) compared intact versus parental remarriage and found the effects on achievement of the former over the latter was d = 0.22, but there were no differences between children from parental remarriage and those children in divorced or widowed families. He argued it is the increased interactions with two adults that is beneficial, but also suggested it may be difficult for children to make more than one family transition (to single or divorced then to remarriage). Kunz (1995) found a d = 0.30 effect, and the effects were slightly lower for school achievement outcomes (d = 0.25 for academic achievement, d = 0.16 for verbal achievement, but d = 0.52 for math achievement). The effects decreased with age, and she related many of the decreases more to the economic differences between one (divorced) and two-parent families. Kunz (2001) was more interested in the effects of divorce had less positive interpersonal relationships with their mother and father, but more positive sibling relationships (although the effects are all very small).

Adopted and non-adopted children

Non-adopted children had slightly higher school achievement than their adopted siblings; the adopted children outperformed their non-adopted siblings and peers who were left behind; adopted students did less well in school than non-adopted children—but the effects are small relative to other influences. Most important, age of adoption seems important.

Those who were adopted in their first year showed no differences (d = 0.09), and the effects increased if they were adopted in the second year (d = 0.32), and beyond their second year (d = 0.42). Although there were fewer studies where it was noted whether the children were subject to abuse, neglect, or malnourishment, the effects demonstrated by these studies were much greater (d = 0.46). Overall, van Ijzendoom and Juffer (2005) concluded that it seemed that "adopted children were able to profit from the positive change of environment offered by adoption and subsequent upbringing" (p. 327), but overall the effects are small.

Only and non-only children

A quantitative review of only-born child literature by Falbo and Polit (1986) found that only-born children surpassed all others except firstborns and children from twochild families concerning achievement and intelligence. In addition only-born children surpassed all non-only children, especially those from families with three or more children in positive character attributes and in the positivity of the parent-child relationship. Only-born children are indistinguishable from firstborns and those from small families across all developmental outcomes. Enhanced parental attention and anxiety are seen as facilitating the development of achievement, intellectual ability, and character. Polit and Falbo (1987) conducted a meta-analysis of the affective differences between onlyborns and other family structures, and identified achievement motivation as the major discriminator (d = 0.17)—only-borns were more motivated and had better relations with parents (d = 0.13)—otherwise there were no differences across many affective outcomes.

Maternal employment

Since the 1980s, the majority of American mothers have been in employment, although the claim that this was somehow detrimental to their children was still a widely held belief. Goldberg, Prause, Lucas-Thompson, and Himsel (2008) showed that the effects of maternal employment on achievement were indeed trivial (r = 0.032). They could not find differences with respect to SES (middle/upper r = -0.043, lower-middle r = -0.055); ethnicity (white r = -0.028, majority African American and Hispanic r = 0.020), child's age (pre-school r = 0.020, elementary r = 0.061, high school r = 0.019), family structure (one parent r = -0.005). It does not matter to a child's achievement whether a mother works outside the home or not.

Home environment

The home environment includes measures of the socio-psychological environment and intellectual stimulation in the home. Iverson and Walberg (1982) suggested that achievement is more closely linked to the socio-psychological environment and intellectual stimulation in the home than to parental socioeconomic status indicators such as occupation and education. They were not specific about which of these home indicators were most influential. Gottfried (1984) completed a meta-analysis on studies using the Home Observation for Measurement of the Environment (HOME) scale, which measures



responsivity, restriction, punishment, play materials, involvement, and variety. The most consistent and highly correlated factors with achievement were maternal involvement, variety, and play materials (Gottfried, 1984).

Television

The overall effects of television on achievement are small but negative; however, given the changes in technologies available to students (video games, computer and interactive technology), the effects of television on achievement are probably of far less interest and importance than most other influences on achievement.

A meta-analysis examining the effect of leisure-time television on school learning (Williams, Haertel, Haertel, & Walberg, 1982) found a small but negative relationship between hours of viewing and achievement. Effects were consistent across sample size, year, and location of the studies. However, the overall effects across the range of viewing times over a week were not constant. There were slightly positive effects for up to ten hours of viewing a week, while over ten hours viewing was related to negative effects with the strength of effects increasing with viewing up to 35 to 40 hours a week. Additional viewing had little effect. This non-linearity in effects is still found in more recent research (Ennemoser & Schneider, 2007). The adverse effects were greater for females and for those with high IQs.

Razel (2001) used six of the larger national and international data bases that asked about television and achievement. The overall effect was negative and there was this same



non-linear effect related to hours of viewing. With up to two hours of television per day, the effects on achievement were small and positive; more than two hours and the achievement to television viewing relationship was negative. As important, the optimal viewing time (that is, when the relationship is at least positively related to achievement) decreases with age—younger children can watch more television with no negative effects, but by age by age seven it is down to one hour, and by 17 it is zero hours.

A similar non-linear effect was reported by Neuman (1988) in a synthesis of findings across eight American states (but not providing an overall effect size). She concluded that there was a curvilinear relation between television viewing and reading skills: children who watched a moderate amount of television (two to three hours daily) scored slightly higher, but those who watched more per day had a slightly lower effect size: But those viewing more than four hours a day had much lower achievement. The variation in effect size between these two groups of viewers, however, was small (d = 0.15). Her argument was that these differences were more reflective of parental characteristics. Parents of children who allowed unrestricted and unsupervised viewing tended to have fewer expectations and lower educational aspirations for their children than those who assumed greater control over television viewing. She found no support for the displacement hypothesis as leisure reading, sports activities, and spending time with friends all seemed independent of the time spent watching television. There can also be positive effects of television on pro-social behaviors (d = 0.63) and this outweighs the effects of anti-social behaviors (d = 0.30, Hearold, 1980).

Parental involvement in learning

There is much variance in the influence of parental involvement. There are negative effects when parents' involvement involves a surveillance approach, lower effects relating to parental involvement in early intervention, and much higher effects relating to parental aspirations and expectations and when parents take a more active approach in learning.

Casto and Lewis (1984) examined studies relating to parental involvement in early intervention programs and found there was little support for the idea that parental involvement leads to more effective intervention programs. They commented that while programs that involved parents could be effective, they were not necessarily more effective than those either not involving parents or involving them in a minor way. Similarly, White, Taylor,



KEY	
Standard error	0.178 (High)
Rank	45th
Number of meta-analyses	11
Number of studies	716
Number of effects	1,783
Number of people (2)	320,000

and Moss (1992) examined the research on parental involvement in early intervention programs and found that claims that parental involvement led to more effective outcomes were without foundation. Often the effect of parental involvement, after the variance due to students and teachers are accounted for, is trivial at best (Innocenti, Huh, & Boyce, 1992).

Hong and Ho (2005) concluded that parent aspirations were the most important influence on their children's achievement, whereas parental supervision in the forms of monitoring students' homework, time watching television, and time going out with friends appeared to have a negative effect on the educational aspirations of adolescent students. Similarly, Rosenzweig (2000) noted that the relationships between student achievement and parental participation (d = 0.56) and supportive parenting (d = 0.43) were much higher than with homework supervision (d = 0.19), participation in school activities (d = 0.14), communication with school and teachers (d = 0.14), monitoring school progress (d = 0.12), providing structure in the home (d = 0.00), and controlling and disciplining parental style (d = -0.09). These effects were the highest in high SES families, in elementary compared to high schools, and in Asian and Latino compared to white and African American families. Of as much interest are those family variables that negatively relate to achievement. These factors included external rewards, homework surveillance, negative control, and restrictions for unsatisfactory grades. Overall, "the higher the hopes and expectations of parents with respect to the educational attainment of their child, the higher the student's own educational expectations and, ultimately, the greater the student's academic achievement" (Hong & Ho, 2005, p. 40). These high expectations are assisted by greater parent-student communication and the student's control over their own studies (see also Fan and Chen, 2001).

Crimm (1992) reviewed parental involvement and found greatest effects between kindergarten and grade 3 (d = 0.41), but these decreased with age (grades 3 to 5 d = 0.36, secondary d = -0.05). The most successful involvement related to tutoring (d = 0.49), and home visits and interactions by teachers (d = 0.48), and the lowest were parent training (d = 0.15). The highest effects were in reading (d = 0.40) while the effects were much lower in mathematics (d = 0.18); which is not too surprising given that parents are more likely to read than do mathematics with their children. Jeynes (2005) found that parental involvement was related to school grades (d = 0.74) and the best predictor was expectations (d = 0.58), which was far greater than parental involvement at the school (d = 0.21). In a subsequent study using secondary students, Jeynes (2007) similarly found



greater effects from parental expectations (d = 0.88) than from other parent factors such as checking homework (d = 0.32), having household rules (d = -0.00), and attendance and participating in school functions (d = 0.14).

Senechal (2006) found that a more active involvement by parents was more effective. For example, the effect size from studies where parents taught their children specific literacy skills were twice as effective (d = 1.15) as parents listening to their children read (d = 0.51), which, in turn, was much more effective than reading to the child (d = 0.18). These effects were reasonably consistent from kindergarten to grade 3, for students with (d = 0.38) and without reading difficulties (d = 0.74), and for families from different socioeconomic status groups.

Home visiting

Sweet and Applebaum (2004) claimed that home visits by school staff not only reduced child abuse but enhanced school achievement. The effect on cognitive outcomes was d = 0.18and on socio-emotional outcomes d = 0.10. Black (1991) was more specifically concerned with the effects of home visiting on learning disabled students. Most visits aimed to offer information and to enhance parental coping and child development; on average there were 36 two-hour visits over a year, The overall effect on cognition outcomes was d = 0.39, the effect on developmental outcomes (birth weight, developmental gains, health status) was d = 0.13, and the effect on social-behavioral outcomes (social functioning, interpersonal, self-esteem) was d = 1.01. This seemed to be a consequence of the more powerful effects of parenting (d = 1.06) and parent social functioning (d = 1.52).

Concluding comments

Parents have major effects in terms of the encouragement and expectations that they transmit to their children. Many parents, however, struggle to comprehend the language of learning and thus are disadvantaged in the methods they use to encourage their children to attain their expectations.

Across all home variables, parental aspirations and expectations for children's educational achievement has the strongest relationship with achievement (d = 0.80), while communication (interest in homework and school work, assistance with homework, discussing school progress: d = 0.38) have a moderate size effect, and parental home supervision (e.g., home rules for watching television, home surroundings conducive to doing school work: d = 0.18) is the weakest. Thus, parents need to hold high aspirations and expectations for their children, and schools need to work in partnership with parents to make their expectations appropriately high and challenging, and then work in partnership with children and the home to realize, and even surpass, these expectations. Too often, the alienation of the home from school reduces the initial expectations. The Flaxmere study, for example, found that, when their children started school, 98 percent of the parents considered that education was very or extremely important to their children's future. Two-thirds of these parents expected their children to attain diplomas and degrees. By the time they left elementary school, these aspirations had been dowsed and the parents mainly wanted their children to "get a job" (Clinton *et al.*, 2007).

Parents should be educated in the language of schooling, so that the home and school can share in the expectations, and the child does not have to live in two worlds—with

horn

little understanding between the home and school. Some parents know how to speak the language of schooling and thus provide an advantage for their children during the school years, while others do not know this language, which can be a major barrier to the home contributing to achievement. Parental expectations are far more powerful than many of the structural factors of the home (e.g., single or two-parent families, families with resident or non-resident fathers, divorced parents, adopted or non-adopted children, or only children and non-only children). It is not so much the structure of the family, but rather the beliefs and expectations of the adults in the home that contributes most to achievement.

Released Under the Official In

The contributions from the school

There have been numerous studies that have attempted to ascertain the amount of variance that can be attributed to the input from the school. Among the most sophisticated are multi-level modeling procedures that can assist in determining this amount of variance relative to other potential influences (Fitz-Gibbon & Kochan, 2000; Teddlie, Reynolds, & Sammons, 2000). This multi-level modeling allows estimation of variability at the student, class, and school levels simultaneously (and assessment of interaction effects across levels). As an example of its use, Konstantopoulos (2005) found that a substantial proportion of the variation in student achievement lies *within* schools and not *between* schools. If the variance is *within*, this means that factors such as teacher variability have a relatively larger effect on student achievement than do school effects. ⁹It appears that the teachers students are assigned to may be more important than the schools they attend" (p. 36).

Alton-Lee (2003) has reviewed many of these studies and ascertained that between zero to 20 percent in student achievement can be attributed to school-level variables and 16 to 60 percent can be attributed to differences between teachers or classes. This spread is critical and seems to be related to specific policies in the various countries from which these data were derived. New Zealand, as an example, has among the lowest percentage of *between*-school variance (about four percent and thus the *within* school variance is much greater.) Using data from the Second International Mathematics Study, Scheerens, Vermeulen, and Pelgrum (1989) found that school effects were undetectable as a source of variance in New Zealand, whereas between-teacher or between-class variance was 42 percent. Harker and Nash (1996; Nash & Harker, 1997) found that the school effect in New Zealand high school performance accounted for between five to ten percent of the variance in mathematics, nine to ten percent of the variance in English, and five to seven percent of the variance in science. The message is that, if you take two students of the same ability, it matters not which school they attend, but it may matter greatly who their teacher is. It is not so much that *teachers matter*, as that the variance within schools indicates that some teachers matter more than others!

These messages about the greater relative importance of teachers than schools are commonplace in this literature. Willms (2000) concluded that "the pressure and support for change needs to be directed at particular teachers within schools, not simply at entire schools" (p. 241, italics in original). Muijs and Reynolds (2001, p. vii) asserted that "all the evidence that has been generated in the school effectiveness research community shows that classrooms are far more important than schools in determining how children perform at school." Rowe and Rowe (1993, p. 15) stated that "on the basis of our findings to date it could be argued that effective schools are only effective to the extent that they

have effective teachers." Bosker and Witziers' (1996) meta-analysis found that about eight percent can be attributed to school-level variance when achievement is adjusted for initial differences between students and schools; and this variance becomes even smaller when other factors are controlled (such as variance between parallel classes and between grades). Scheerens and Bosker (1997) found that, when adjusting for initial differences between students, schools account for eight percent of the achievement differences.

The situation is quite different in less resourced nations (e.g., throughout Africa) where most variability is *between* schools (Bosker & Witziers, 1996); and in countries where there are high levels of stratification in school types (e.g., academic and vocational). Similarly the teacher variance is lower for achievement in elementary school reading where family and community input is comparably strong whereas teacher variance tends to be higher for mathematics and other curriculum areas that are less directly linked to everyday experiences of students (especially in the home).

One of the sobering conclusions from the above summary is that many of the influences that really make a difference to student learning in developed nations are within schools, from the influence of specific teachers, specific curriculum, and strategies teachers use to teach (Grodsky & Gamoran, 2003). Another important consideration is that there are many more factors *within* schools than teachers—here are also the effects of the culture and ethos of schools, the effects of the principal, and class compositional effects.

This chapter is divided into six major sections:

- 1 attributes of schools (e.g., finances, types of schools);
- 2 school compositional effects (e.g., school size, mobility, mainstreaming);
- 3 leadership;
- 4 classroom compositional effects (e.g., class size, ability grouping, retention);
- 5 school curriculum effects (e.g., acceleration, enrichment);
- 6 classroom influences (e.g., climate, peer influences, disruptive behavior).

Attributes of schools

Finances

Although the meta-analyses seem to indicate that money does not matter, this would be a misleading conclusion. Childs and Shakeshaft (1986) undertook a meta-analysis of studies



KEY	
Standard error	na
Rank	99th
Number of meta-analyses	4
Number of studies	189
Number of effects	681
Number of people (2)	2,277,017

Table 6.1 Summary information from the meta-analyses on the contributions from the school

School	No. metas	No. studies	No. people	No. effects	d	SE	CLE	Rank
Attributes of schools								
School effects	I	168		168	0.48	0.019	34%	50
Finances	4	189	2,277,017	681	0.23		16%	99
Types of school								
Charter schools	I.	18		18	0.20	_	14%	107
Religious schools	2	71	54,060	71	0.23		16%	101
Summer schools	3	105	28,700	600	0.23		16%	98
Desegregation	10	335	6,73 I	723	0.28	0.060	20%	91
College halls of residence	I	10	11,581	23	0.05	_	3%	130
School composition effects							\times	
School size	I	21	—	120	0.43	-0	30%	59
Summer vacation	I	39	—	62	-0.09	<u> </u>	-6%	134
Mobility	3	181	185,635	540	-0.34	0.005	-24%	138
Out of school experiences	2	52	30,554	50	0.09	—	6%	127
Principals/school leaders	11	491	1,133,657	1,257	0.36	0.03	25%	74
Classroom composition effects				-	\sim			
Class size	3	96	550,339	785	0.21	_	15%	106
Open vs tradiitonal	4	315	—	333	0.01	0.032	0%	133
Ability grouping	14	500		1,369	0.12	0.045	9 %	121
Multi-grade/age classes	3	94	<u> </u>	72	0.04	_	3%	131
Within-class grouping	2	129	16,073	181	0.16	—	11%	116
Small group learning	2	78	3,472	155	0.49	_	34%	48
Mainstreaming	5	150	29,532	370	0.28	_	19%	92
Retention	7	207	13,938	2,675	-0.16		-11%	136
Curricula for gifted students								
Ability grouping for gifted	5	125		202	0.30	0.064	21%	87
students								
Acceleration	2	37	4,340	24	0.88	0.183	62%	5
Enrichment	🧹 3 🎽	214	36,336	543	0.39	0.018	28%	68
Classroom influences								
Classroom management		100		5	0.52	—	37%	42
Classroom cohesion	3	88	26,507	841	0.53	0.016	38%	39
Classroom behavioral	3	160	0	942	0.80	0.290	56%	6
Decreasing disruptive behavior	3	165	8,426	416	0.34	0.037	24%	80
Peer influences	I	12	—	122	0.53		37%	41
Total	101	4,150	4,416,898	13,348	0.23	0.072	16%	_

on the relationship between educational expenditure and student achievement and showed that there was a minimal relationship between the two, and the most positive relationship related directly to the costs of instruction; for example for teacher salaries and instructional supplies. Teacher salaries, in turn, were more related to years of teaching experience and not teacher quality. Rolle (2004) also argued that more money was not necessarily needed but that there should be more productive use of existing resources. This is consistent with the claims often made by Hanushek (1989) that there is no consistent statistical relation between educational expenditure and measures of student performance. For example, Hanushek (2003) correlated high school spending per pupil and mathematics scores across 23 countries (from TIMSS, 1998) and found a correlation of r = 0.06. So often money is

added into the education system with little attention to the efficiency or effectiveness of education outcomes. It is not the amount of money spent that is important, but how it is spent.

Murdock (1987) reported financial aid on college students had a small but positive effect on student persistence, and enabled low-income students to persist at a similar rate to that of students from other socioeconomic groups. While student financial aid is an important tool in helping students to stay in college, other factors such as the type of institution, the length of course, and year in which aid is provided (effects are greater with more senior students than with first-year students) all mediate the effects of financial aid.

In a rebuttal to the claims about the limited effect of increased finances, Hedges, Laine and Greenwald (1994; Greenwald, Hedges, & Laine, 1996) analyzed the effects of differential school inputs on student outcomes. Their analysis showed systematic, positive patterns in the relations between educational resource inputs and student outcomes. An increase in per pupil expenditure of \$500 increased the effect on achievement by d = 0.15 for perpupil expenditure, d = 0.22 for teacher education, d = 0.18 for teacher experience, d = 0.16 for teacher salary, and d = 0.04 for teacher/student ratio. Thus we can expect "comparable and substantial increases in achievement if resources were targeted to selecting (or retaining) more educated or more experienced teachers" (Greenwald *et al.*, 1996, p. 380). There is little evidence, however, to justify the notion of "substantial", but there is much consistency with other meta-analyses about the importance of the teacher (and costs associated with enhancing teaching).

The seemingly limited effect on finances can be related to the fact that (a) most studies have been conducted in well resourced countries (e.g., the United States, the United Kingdom) where the variance in resources to schools is not so substantial; (b) most finances in schools are tied up not in discretionary but in fixed costs (such as teacher's salaries, busing, and buildings) and these do not vary in proportion of costs across schools within any country; and (c) if the school composition effects are much greater at the within- than between-school level, then costs could make a difference within schools and less of a difference between schools (and most are currently focused on between- and not within- schools). As Hanushek (2002; 2003) has long argued, there are few incentives for a teacher to maximize achievement, as most of the financial incentives are related more directly to school rather than teacher differences.

The emphasis may need to be *not* on the notion of "Does money make a difference?" but on "*How* does money make a difference, particularly above and beyond the fixed costs of running a school (capital, lighting, salaries)?" It is difficult to imagine money does not make a difference at these critical margins. Jonathon Kozol (2005), in his scathing analysis of the restoration of apartheid schooling in America, *The Shame of the Nation*, cites Deborah Meier's comment that "I'll believe money doesn't count the day the rich stop spending so much money on their own children."

Types of schools

Charter schools

Charter schools have been one of the fastest growing sectors in the United States, and are often aimed to provide, so claim the proponents, what the public schools cannot. Charter schools are publicly funded schools that have been freed from some of the regulations and



statutes that apply to other public schools. They are often set up as autonomous schools competing with public schools, by non-profit groups or universities, often with a particular flavor, and they usually involve some form of innovative teaching principles. In return, they are expected to have high levels of accountability for student outcomes.

Miron and Nelson (2001) found an effect size of d = 0.20 when comparing achievement in charter and regular schools, but when the lower quality studies were excluded, this difference dropped to zero. They concluded that, in spite of the topic's importance to the debate over charter schools and school reform, it is striking how little we currently know about the effect of charter schools on student achievement. They noted that only eight of the 38 states with charter school laws had useable independent evaluations of achievement effects. Not surprisingly, given the close to zero effect, there is a mixture of positive and negative effects, and there is much variation across the states. The hype and promise is much greater than the effects on student achievement.

Religious schools

As was the case with charter schools, much has been written about how different religious schools are from public schools—and indeed they should be. Many have claimed that students in religious schools outperform their public school peers, mainly because of the increased attention to the teacher-student relationship, a greater fostering of parent-school interactions, shared values between families and school, the underlying philosophies of caring and commitment, and a higher work ethic (e.g., Coleman, 1992;



Russo & Rogus, 1998). The dividend of attending religious schools is supposedly greater for those from lower socioeconomic backgrounds. Jeynes (2002) found that African American and Hispanic students attending religious schools (mainly Christian schools) had a d = 0.25 effect size increase over those who attended a public school. The effects were similar in reading (d = 0.25) and mathematics (d = 0.25), and also higher on school-related behavior (d = 0.32). These effects could not be attributed to differing socioeconomic variables, were remarkably consistent in favor of religious schools, and slightly stronger for high school and middle school than elementary school students.

Summer schools

Does going to summer school make a difference? In general, not much, but it is difficult to ignore even these small gains if they are critical to students who may be already marginal (as that is often the criteria for selection). Cooper, Charlton, Valentine, Muhlenbruck, and Borman (2000) analyzed 93 summer programs and their students scored about d = 0.23 greater than those not in summer schools, although the effects were more positive for middle-class than students from disadvantaged backgrounds. Higher effects were found for programs more specifically tailored to the student needs, when parents were involved, for mathematics more than reading, and the effects were the same across all grade levels. Both Cooper *et al.* (2000) and Kim (2002) also found small effects from both remedial and acceleration summer programs.

This pattern was replicated by Kim (2002). He found no significant differences relating to the purposes of the summer school: remediation d = 0.16, enrichment d = 0.16, bridging to high school d = 0.25, assisting grade promotion d = 0.21. Nor were there differences as to whether the summer school was related (d = 0.22) or not (d = 0.14) to the school curriculum, whether it was monitored (d = 0.16) or not (d = 0.16), whether there was a teacher training component for the summer school (d = 0.21) or not (d = 0.14), whether it used current teachers at the school (d = 0.12) or not (d = 0.17) whether food was offered (d = 0.18) or not (d = 0.14), whether the class sizes were above 25 (d = 0.18) or below 25 (d = 0.15), or whether the number of hours of instruction were high (< 132 hours d = 0.11) or not (d = 0.12), and middle and higher SES students (d = 0.21) gained more than lower SES students (d = 0.12).



KEY	
Standard error	na
Rank	98th
Number of meta-analyses	3
Number of studies	105
Number of effects	600
Number of people (2)	28,700

Desegregation

Desegregation is a process aimed at ending racial segregation, typically used in reference to the United States, and started in earnest following the Civil Rights legislation in the 1960s. It was argued that by such measures the United States would be more likely to achieve the more ambitious goal of racial integration. McEvoy's (1982) meta-analysis comparing the effect of desegregation on African American students in desegregated and non-desegregated groups concluded that students in desegregated schools performed at higher achievement levels than students in control groups (d = 0.20), although he reported no differences on self-esteem. He found that studies with control groups showed greater effect than studies without controls (d = 0.48 compared to -0.09); and studies of more than one year demonstrated greater effect than those of less than one year (d = 0.27 compared to -0.07). There was not that much difference between mathematics and verbal language skills (d = 0.28 and d = 0.20), or between elementary and high school students (d = 0.22 and d = 0.18, respectively).

In contrast to findings on the positive effects of desegregation, both Armor (1983) and Krol (1980) found there to be virtually no effects from desegregation on the achievement of African American students in reading and mathematics. Crain and Mahard (1982) also found that desegregation improved achievement for African American students by about d = 0.08, but noted that there were only marked effects in the earliest primary grades (d = 0.44). Two methodological factors correlated with the measured effect of desegregation on academic achievement: studies where students received only partial treatment (i.e., began desegregation after completing one or more years of segregated schooling) and the type of control group. Stephan (1983) found that desegregation resulted in improvement in the reading achievement results of African American students but there was no effect on mathematics. Younger students benefited more than older students in reading. In a small number of studies where desegregation was voluntary, reading achievement was significantly better; however the number of studies means this finding is not conclusive. Miller and Carlson (1982) also noted that while they found desegregation to have had a moderate positive effect on the academic achievement of African American students, there was improvement in verbal but not mathematics achievement.

Wortman and Bryant (1985) analyzed much the same data as Stephan (1983) and Krol (1980)—although Wortman and Bryant rejected articles they considered of lower quality, thus removing 79 articles. They then reported a mean effect of d = 0.45 but noted that effects "for the better designed quasi-experiments are considerably smaller" (Wortman & Bryant, 1985, p. 304)—reducing the mean effect to d = 0.20 for those with no selection



KEY	
Standard error	0.060 (Medium)
Rank	91st
Number of meta-anal	yses 10
Number of studies	335
Number of effects	723
Number of people (1)	6,731

problems. They found similar effects for elementary (d = 0.43) and high school (d = 0.55), and reading (d = 0.57) was higher than mathematics (d = 0.33).

In a different ethnic comparison, Goldring and Addi (1989), examining the ethnic composition of the classroom and in reading comprehension achievement in Israel, found that integrated classrooms, compared to minority segregated classrooms, provided a better learning environment for students of both Asian-African and western origins.

Overall, desegregation is a topic where the meta-analyses show a wide variation in the effect sizes typically relating to the selection process for the inclusion of articles. It is likely that there are many more critical factors than the composition of the classroom that affect achievement—and the success of desegregation may be better assessed to the degree it provides opportunities and diversity than achievement effects.

College halls of residence

Interest in whether residing in residential halls or not has an affect on achievement has been primarily the domain of colleges and universities. Blimling (1999) found that it did not matter whether a student lived in a college, at home, in a fraternity or sorority house, or in off-campus housing or flats. His message was that the zero effect he found should lead to many institutions (such as residential halls) seriously questioning the educational value they were adding to student learning—and clearly they are not adding value at the moment.

School compositional effects

School compositional effects include the size of schools, the effects of summer vacation, mobility, and out-of-school experiences.

School size

Another school level effect is the enrollment size of the school. Stekelenburg (1991) found effects of d = 0.47 between size of high school and achievement, which is quite substantial for a structural effect, although he considered these relatively small. He argued that while very small schools can be expensive to operate, the curriculum advantages of larger schools start to reduce in their effectiveness as they grow much beyond 800. He considered the optimal size to be about 800, and argued that the "smaller the high school, the more it



KEY	
Standard error	na
Rank	130th
Number of meta-analyses	1
Number of studies	10
Number of effects	23
Number of people (1)	11,581