The Springer Series on Human Exceptionality

Anastasiya A. Lipnevich
Franzis Preckel
Richard D. Roberts
Editors

Psychosocial Skills and School Systems in the 21st Century

Theory, Research, and Practice

Springer

The Springer Series on Human Exceptionality

More information about this series at http://www.springer.com/series/6450

Anastasiya A. Lipnevich • Franzis Preckel
Richard D. Roberts
Editors

Psychosocial Skills and School Systems in the 21st Century

Theory, Research, and Practice

Springer

Editors
Anastasiya A. Lipnevich
Educational Psychology
Queens College and the Graduate Center
The City University of New York
New York, NY, USA

Franzis Preckel
Department of Psychology
University of Trier
Trier, Germany

Richard D. Roberts
Center for Innovative Assessments
Professional Examination Service
New York, NY, USA

ISSN 1572-5642
The Springer Series on Human Exceptionality
ISBN 978-3-319-28604-4 ISBN 978-3-319-28606-8 (eBook)
DOI 10.1007/978-3-319-28606-8

Library of Congress Control Number: 2016931335

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Foreword

Remember the "Dead Poets Society"? An inspiring combination of wit, intelligence, motivation, and tragedy. Neil, Todd, Knox, and Charlie. Such a memorable amalgamation of similarities and differences, such a drama of intelligence and emotions. Such different outcomes.

It is quite remarkable that the last two decades of studies performed at the crossroads of psychology and education have been dominated by constructs similar to those captured in so many stories depicted in the "Dead Poets Society," many other movies and books, and in the story about Amanda and Rachel (see Hulleman, Barron, Kosovich, & Lazowski, Chap. 10). This is not to say that the traditional notion of abilities, as captured by the g-factor and other variations on that theme, have disappeared. Not at all. Those studies are still alive and well. Yet, given that the conventional ability-driven line of research in psychology has had a much longer history, the parameters of the relevant knowledge (e.g., the predictive power of traditional ability-related concepts for educational outcomes) are well known and, correspondingly, do not generate much curiosity and excitement in researchers. The fashion today is to look at "everything else" that is not captured by traditional definitions of ability. That everything else may be referred to by many different words, among which are executive functioning, noncognitive factors, learning styles, personality, emotions, and various combinations of these constructs, but the bottom line is that the era of conventional ability-based predictions of school achievement is over. The King is dead, long live the King! This new King is the central figure of the discourse presented in this volume edited by Anastasiya A. Lipnevich, Franzis Preckel, and Richard D. Roberts.

This volume is important for a number of different reasons. First, it brilliantly captures the sheer breadth and width of this new Kingdom. By definition, the Kingdom has merged peaceably a number of academic domains of psychology – social, personality, cognitive, educational, and organizational and management. Correspondingly, it has been published on by journals and books in all of these different domains as well as by interdisciplinary outlets. For example, the journal that

I edit, *Learning and Individual Differences*, received more than 500 submissions in 2015, and about 75 % of them were representative of the work in the Kingdom. If this figure is multiplied by the dozens and dozens of journals that publish on such research, one can easily appreciate the scope of the field.

Second, this volume nicely reflects an array of concepts, definitions, and approaches exercised in the Kingdom. It also provides examples of careful cross-mappings of the elements of this array, as exemplified in Chapter 1 by Jeremy Burrus and Meghan Brenneman (see especially, Table 1) especially Table 1 of the chapter. Clearly, given the huge conceptual landscape that is engaged even in this single volume, more of such careful vocabulary work is needed to figure out how, if at all, grit differs from motivational determination, on the one hand, and thought perseverance on the other.

Third, gently, but clearly, the book brings up a very important point as it relates to the relevant constructs: that of the quantity and quality of the assessments and measurement devices used in the Kingdom (Matthias Ziegler and Martin Brunner, Chapter 2). Indeed, there is a danger in working with assessment devices that appear to be easy to construct and, thus, are introduced to the field without the proper foundational work.

Fourth, as reflected by the volume, today, the crossroads between psychology and education are overrun by intervention work. The mosaic of intervention programs that are in action out there is quite breathtaking. This volume sets the stage for the much-needed careful analyses of effectiveness, cost-benefit figures, and values that are imposed on children through these programs.

Fifth, the book raises the question of the possible policy-relevant interpretations and most effective applications of the work. What should be done with these findings? Should not-ability-like (i.e., captured by means of "soft" rather than "maximum-performance" assessments) indicators be used in situations of high-stakes decision-making? What is the place of the relevant constructs in various situations charged with consequences?

All in all, the volume provides an engaging read and is a colorful quilt of ideas, approaches, and interpretations. It has numerous take-home messages that are worth the reader's attention, but the main one is that the dominance of conventional abilities at the crossroads of education and psychology is over. The King is dead, long live the King!

Elena L. Grigorenko

Preface

The focus of this edited volume is on constructs that are commonly referred to as psychosocial skills in the education and psychology literatures. These characteristics have countless alternative names in economics, policy, and other closely related fields, including noncognitive characteristics, socioemotional skills, soft skills, affective skills, personal skills, dispositions, twenty-first-century skills, character, and many others. The unifying theme for this plethora of synonyms and name variants is in their differentiation from cognitive skills, intelligence, or so-called hard skills.

Since the very inception of formalized education there has been recognition of the key importance of psychosocial skills and the pivotal role they play in the academic and life success of an individual. Confucius, Aristotle, Jean-Jaques Rousseau, Immanuel Kant, John Dewey, Martin Luther King – to name a few – all acknowledged that character development should constitute one of the focal points of the educational process. Parents, teachers, and athletic and youth club leaders recognized the inherent value of such skills and their contribution to the enhancement of cognitive skills, but for a long time such recognition remained implicit and detached from a formally presented instructional curriculum. Thus, until recently, the majority of academic programs have been primarily focused on knowledge acquisition and enhancement of cognitive skills related to the three Rs, or *r*eading, w*r*iting, and the a*r*ithmetic. This picture, however, is changing with psychosocial skills and their systematic development been inculcated into policy as an inherent task of educational institutions. Indeed, a strong commitment to developing psychosocial skills can be found in the mission statements of many schools across the globe (e.g., Stemler & Bebell, 2012), and in national policy statements worldwide.

One of the main goals of any education system is to prepare individuals for a successful and fulfilling life by assisting them in the development of a range of skills that would be necessary for finding a rewarding occupation in the future. A report, entitled "Are They Really Ready to Work?," asked employers from various fields to list the most valued skills that ensure success in the workplace. Work ethic, teamwork, oral communication, leadership, creativity, and lifelong learning topped

the list of characteristics held as important by employers (Casner-Lotto & Barrington, 2006). Cognitive skills were deemed as critical also but second to psychosocial characteristics. Most importantly, the report revealed that psychosocial skills were precisely the skills most often found lacking in new employees joining the workforce (i.e., former students of the K-12 system).

So what are the skills that matter most in the context of K-12 education? The Partnership for 21st Century Skills (P21) has started its work with a mile-long list of the skills considered necessary for success in the twenty-first century. Then, in their strive for parsimony, P21 condensed the long list of skills to four, which they presently call the 4 Cs: critical thinking, communication, collaboration, and creativity. A fifth skill, self-regulation, also frequently appears in their publications. Another advocate for the importance of psychosocial skills, the Collaborative for Academic, Social, and Emotional Learning (CASEL) arrived at five "competency clusters," namely self-awareness, self-management, social awareness, relationship skills, and responsible decision-making. A number of bestsellers that tackle the topic of the importance of psychosocial skills present their own perspective on what constitutes the core of critical psychosocial skills. Paul Tough's (2013) *How Children Succeed* lists grit, curiosity, and the hidden power of character to be of utmost value. Madeline Levine (2013) in her *Teach Your Children Well: Why Values and Coping Skills Matter More Than Grades, Trophies, or "Fat Envelopes"* emphasizes the importance of resilience, coping skills, and intrinsic motivation. Tony Wagner (2010), in turn, in his *The Global Achievement Gap* names the "seven survival skills" that include problem-solving and critical thinking, collaboration across networks and leading by influence, agility and adaptability, initiative and entrepreneurship, effective written and oral communication, accessing and analyzing information, and curiosity and imagination. The examples are plentiful, and there are more taxonomies, models, and lists of key skills, developed and used by practitioners, policymakers, and researchers that we have the space to cover in this preface.

Indeed, the task of finding the most meaningful skills that should be the focus of everyone's attention is daunting. Matters are even complicated by an inflation of "new" skills, many of which appear upon closer inspection to be "old wine in new bottles". Understanding how central skills develop, what are the best ways to measure them, and how to cultivate them in students – are all questions waiting to be answered, or at least systematized. Hence, the main purpose of this edited volume is to present an overview of the extant literature on psychosocial skills in education and to offer a general conceptual framework in which the wealth of constructs discussed as psychosocial skills can be integrated. The book will paint a comprehensive picture of the current state of research on psychosocial skills, discussing the scientific status, theory, assessment, developmental trajectories, and areas of application of specific constructs that have been deemed critical to educational and life success. The reader will become acquainted with the questions that are being asked and the issues that researchers and practitioners face and will learn about the methodologies and theoretical perspectives that are used to address them. Table 1 presents the list of psychosocial skills discussed in the current volume.

Table 1 Psychosocial constructs discusses in the current volume

Construct	Chapter
Conscientiousness	Chapters 1, 4, 7
Coping	Chapters 1, 11
Creativity	Chapters 1, 4, 6
Curiosity	Chapters 1, 5
Emotional intelligence	Chapters 3, 11, 13
Emotional stability	Chapters 1, 4, 11
Empathy	Chapter 3
Extraversion	Chapters 1
Grit	Chapters 1, 7, 14
Intellectual engagement	Chapters 1, 5
Intrinsic motivation	Chapters 1, 3, 6, 7, 8, 10, 11, 14
Leadership	Chapters 1, 3
Need for cognition	Chapters 1, 5
Openness to experience	Chapters 1, 4, 5
Perfectionism	Chapters 1, 7
Resiliency	Chapter 12
Self-concept	Chapters 8, 10
Self-directed learning	Chapter 3
Self-efficacy	Chapters 1, 5, 9, 10
Self-esteem	Chapters 1, 3, 10, 12
Self-regulated learning	Chapters 9, 10, 11
Teamwork	Chapters 1, 3
Time management	Chapters 1, 7

The volume includes four main sections. Part I entitled, *General Background: Theory and Guiding Principles*, focuses on existing conceptual models and psychometric considerations, before moving to consider approaches to the assessment of psychosocial skills and how these develop over time. Burrus and Brenneman (Chap. 1) open this section by presenting an organizational framework for psychosocial skills. The proposed framework draws upon a prominent taxonomy of personality, the Five Factor Model. These contributors begin by mapping existing psychosocial skills onto the Five Factor Model, before further subdividing these constructs into three categories of skills that appear important for K-12 students: *performance skills* ("getting along with school"), *interpersonal skills* ("getting along with others"), and *self-management skills* ("getting along with yourself"). In so doing, Burrus and Brenneman offer a basic orientation to the reader for organizing the various skills found throughout the remainder of the book. Chapter 2 deals with psychometric assessment and its tenants as a major cornerstone of educational science and practice. Brunner and Ziegler discuss key test standards that are relevant when constructing or selecting tests for psychosocial assessments in education. The researchers present basic principles that are associated with test score reliability and the validity of test score interpretations, as well as psychometric models and their

importance for computing and understanding the veracity of psychosocial test scores. Chapter 3 of this section, by Stemler and DePascale, identifies key skills that appear in mission statements of schools and intends to provide school leaders and policymakers with a reference that will help them to easily identify strong psychometric measures of the skills and competencies that they aim to foster in their students. Part I concludes with Walton and Billera (Chap. 4), who review the current state of research on personality development across the life span (with special emphasis on childhood and adolescence). While historically, childhood temperament was studied in isolation from adult personality Walton and Billera use the Five Factor Model as a feasible framework for presenting stability and change in personality from childhood to adulthood. Notably, the evidence presented here also establishes an empirical basis for considering the psychosocial skill interventions covered later in the book highly feasible.

The second section, entitled *Psychosocial Skills: Key Constructs*, deals with specific constructs deemed as critical in the K-12 context. Following the call of Burrus and Brenneman (Chap. 1), the chapters in this section can be readily (though not entirely perfectly) mapped onto the Big Five framework. As such, we use this as an organizing principle for these chapters. The Openness factor of the Big Five aligns closely with both Need for Cognition (Jebb, Saef, Parrigon, and Woo, Chap. 5) and Creativity (Kaufman, Beghetto, and Dilley, Chap. 6). Contributors in these two chapters provide extensive reviews of the state of the art in these two domains. By contrast, Kim, Poropat, and MacCann (Chap. 7) present a review of the Conscientiousness factor, which arguably has some of the most impressive validity evidence for subsequent educational policy and practice. By contrast to rather focused assessment on factors comprising the Big Five, Trautwein and Möller (Chap. 8) discuss the issue of self-concept, Bembenutti, White, and DiBenedetto (Chap. 9) focus on self-regulated learning, and Hulleman, Barron, Kosovich, and Lazowski (Chap. 10) explicate the concept of student motivation. These three chapters cut across the Big Five Factors of Openness, Conscientiousness, and Emotional Stability. Chapter 11 by Goetz and Bieg concludes this section by discussing academic emotions and emotional intelligence, which generally represent a construct that cuts across the Emotional Stability and Agreeableness factors of the Big Five (as well as invoking some components of cognitive skills). Each chapter in Part II includes definitions of the highlighted construct, relationships with key educational outcomes, and reviews intervention programs geared to enhance each psychosocial characteristic.

The third section of the volume, *Psychosocial Skills: Applications*, focuses on educational interventions, describing the place of psychosocial skills in large-scale international assessments, policy and legislature, as well as their current place in curriculum design and development. Prince-Embury, Keefer, and Saklofske (Chap. 12) review research on the development of academic resilience. These researchers review theory and provide examples of applications of resiliency and related strength-based constructs in schools at different levels of intervention: school-wide/systemic, classroom, and individual. Chapter 13 by Torrente, Rivers, and Brackett offers an overview of findings related to the development of emotional intelligence. More

specifically, the chapter focuses on the RULER program to illustrate how the teaching and learning of emotional intelligence can be integrated into core academic curricula. Finally, Bertling, Borgonovi, and Almonte (Chap. 14) cover historic and current trends on the role of psychosocial skills in large-scale group-score assessments.

In the final section of this book, *Conclusions*, the editors (Lipnevich, Preckel, & Roberts, Chap. 15) join forces in an attempt to integrate these various elements into a coherent whole. In the process, we make a series of recommendations for future research, policy, and practice. A true labor of love, we hope you like it!

All in all, this book provides a comprehensive summary and evaluation of current research and may serve as a resource for practitioners who might now make informed decisions on which psychosocial skills to focus upon in their respective educational programs. Humbly, we hope and trust that researchers, students, policymakers, and practitioners alike will find this volume to be of some value.

New York, NY, USA — Anastasiya A. Lipnevich
Trier, Germany — Franzis Preckel
New York, NY, USA — Richard D. Roberts

Acknowledgments

We would like to thank our esteemed colleague Professor Don Saklofske for inviting us to join in on this effort and prepare a volume for *The Springer Series on Human Exceptionality*, which he edits together with Moshe Zeidner. The idea to write a book that presents a comprehensive review of psychosocial skills research and offers a mixture of theory, assessment, development, and applications has been evolving for many years. We are indebted to Don for offering us the opportunity to create, compile, and finally deliver this book to you, the reader. We would like to also thank Professor Elena Grigorenko, a recognized authority in the field of both cognitive and psychosocial skills (and so much more), who graciously agreed to write the Foreword to this edited volume.

During the writing of our respective chapters and the editing of this book, Queens College and the Graduate Center of the City University of New York, the University of Trier, and Professional Examination Service provided the facilities and resources necessary to undertake and complete this work. Being a book on psychosocial skills, and the importance of demonstrating appropriate behavior in context, we would be remiss not to extend our gratitude to senior management of these respective institutions, most notably Robert Block (CEO) and Simmy Ziv-el (Vice-President) at Professional Examination Service.

Edited volumes can be very difficult to produce, as anyone who has set out on this onerous task will testify. We are indebted to Dana Marie Murano, Christina Cappiello, and Alicia Ogilvie, for their work pulling together the important preproduction pieces, including formatting all the chapters, compiling tables, and arranging many meetings we had as the editorial team. Special thanks to the production team at Springer for making the final stretch of this process tolerable – Judy Jones, Michelle Tam, and Kamalavasanthi were fantastic at pushing, encouraging, and cajoling us along the way, and answering our countless questions with no complaints whatsoever.

No arduous undertaking can reach completion without continuous support, steadfast encouragement, and difficult sacrifice from all involved. First and foremost, we would like to thank our wonderful children – Emilian and Evan Lipnevich, Tanno and Valerian, and Matthew Dean Roberts and Caspian Sondre Aicher-

Roberts – who have variously lost us for stretches of time as we dealt with logistic issues, reviewed one of the contributions by our distinguished panel, and wrote our own pieces. We are blessed to have you in our lives and hope and trust that as you move along your academic paths our work on this volume and beyond, directly or indirectly, will be of help to you. We love you dearly.

Many thanks to our partners, Sergey Lipnevich and Bernd, for helping us through this endeavor. You steadily believed, constantly reassured, relentlessly cheered, and persistently inspired, while taking care of everyday issues, to which we could not attend. We could not have done it without you. And to the following friends, family, students, and/or colleagues, many thanks to you for your advice, for listening to us as we vented our frustrations, and for just plain being there: Larisa, Andrey, and Viktoria Tsobkalo, Elena and Stepan Khinevich, Jeffrey and Lisa Smith, Maksim Rutenberg, Stefan Krumm, Ludmila Lubane, Kalina Gjicali, Eva Ostrum, Julia White, Diana Krivtsova, Nicola Baumann, Michael Schneider, Thomas Götz, Trent Raznor, Rebecca Rhodes, Carolyn MacCann, Jeremy Burrus, Gabriel Razvan Olaru, Ralf Schulze, Simmy "General Groves" Ziv-el, Oliver Wilhelm, Rob Schneider, Gerald Matthews, Selina Weiss, and Franklin Zaromb.

Finally, an edited book would be nothing without its contributions from the acclaimed scholars who we asked to give of their time, expertise, and knowledge to prepare chapters on topics that we felt needed special attention. We are indebted to each and every one of you for the various chapters appearing in this volume. We appreciate your critical contributions, willingness to cope with a challenging task, and to respond to our suggestions with grace, dignity, and efficiency. Because of your joint efforts, we believe the volume is coherent and hugely informative.

We hope this book will give readers a deeper understanding and appreciation of the current state of the art in psychosocial skills research and how it operates in K-12 settings. Please enjoy this volume and do not hesitate to drop us a line should the book raise some questions.

Anastasiya A. Lipnevich
Franzis Preckel
Richard D. Roberts

Contents

About the Editors

Anastasiya A. Lipnevich, Ph.D., is an Associate Professor of Educational Psychology at Queens College and the Graduate Center of the City University of New York. Her research interests include instructional feedback, attitudes toward mathematics, alternative approaches to non-cognitive assessment, academic emotions, and the role of psychosocial characteristics in individuals' academic and life achievement. Dr. Lipnevich has over 40 publications in refereed journals and edited volumes, has delivered more than 60 conference presentations and invited addresses, has chaired 10 symposia, serves on several editorial boards, and has co-edited several special issues in flagship journals. Dr. Lipnevich's professional honors include, among others, Excellence in Dissertation Award from Rutgers University, the New Investigator Award 2010 from Division 3 (Experimental Psychology) of the American Psychological Association, and a Postdoctoral Fellowship from Educational Testing Service.

Franzis Preckel, Ph.D., is Professor for Giftedness Research and Education at the Department of Psychology of the University of Trier, Germany. Her main research interests are intelligence (assessment and structural analysis), giftedness (fostering options and program evaluation), cognitive motivation (development and predictive validity for academic achievement), self-concept (development and reference group effects), and the influence of chronotypes on student learning and achievement. Dr. Preckel has edited or coauthored 19 books and published 46 articles in peer-refereed journals, as well as 15 book chapters. She is coauthor of 5 intelligence tests and 2 self-report scales on academic anxiety and on cognitive motivation. She has presented her research in more than 80 talks and is on the editorial board of Diagnostica and Gifted and Talented International.

Richard D. Roberts, Ph.D., is Vice President and Chief Scientist, Center for Innovative Assessments, Professional Examination Service, New York, New York. A former National Research Council Fellow who conducted research at Brooks Air Force Base (1996–1998), he has also been a Senior Lecturer at The University of Sydney (1998–2003), and a Managing Principal Research Scientist at the Educational Testing Service (2003–2014). His main area of specialization is measurement, with a special emphasis on developing and researching innovative new

items types for the assessment of both cognitive and noncognitive factors, as well as constructs that have elements that straddle these two domains (e.g., emotional intelligence). Dr. Roberts has co-athored over a dozen books, and about 200 peer-review articles and book chapters on these topics in diverse sub-disciplines (including, education, psychology, business, medicine, health policy, and wind engineering), with nearly 400 presentations across the globe. Among Dr. Roberts' professional honors are two ETS Presidential Awards, two PROSE book awards, and various early (and then later) career awards. He is currently Secretary-Treasurer of the International Society for the Study of Individual Differences and serves on the boards of various international journals and organizations.

Contributors

Debby E. Almonte Educational Testing Service, Princeton, USA

Kenneth E. Barron Department of Psychology, James Madison University, Harrisonburg, VA, USA

Ronald A. Beghetto Educational Psychology, Neag School of Education, University of Connecticut, Storrs, CT, USA

Héfer Bembenutty Department of Secondary Education and Youth Services, Queens College of the City University of New York, Queens, NY, USA

Jonas P. Bertling Educational Testing Service, Princeton, USA

Madeleine Bieg Empirical Educational Research, University of Konstanz, Konstanz, Germany

Kimberly A. Billera Department of Psychology, St. John's University, Jamaica, NY, USA

Francesca Borgonovi OECD, Paris, France

Marc A. Brackett Yale Center for Emotional Intelligence, Yale University, New Haven, CT, USA

Meghan Brenneman Research and Development, Educational Testing Service, Princeton, NJ, USA

Martin Brunner Department of Education and Psychology, Free University of Berlin, Berlin, Germany

Jeremy Burrus Center for Innovative Assessments, Professional Examination Service, New York, NY, USA

Mary DePascale Department of Psychology, Wesleyan University, Middletown, CT, USA

Maria K. DiBenedetto Bishop McGuinness Catholic High School, Kernersville, NC, USA

Anna Dilley Educational Psychology, Neag School of Education, University of Connecticut, Storrs, CT, USA

Thomas Goetz Empirical Educational Research, University of Konstanz, Konstanz, Germany

Thurgau University of Teacher Education, Kreuzlingen, Switzerland

Chris S. Hulleman Department of Educational Leadership, Policy, and Foundations, University of Virginia, Charlottesville, VA, USA

Andrew T. Jebb Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA

James C. Kaufman Educational Psychology, Neag School of Education, University of Connecticut, Storrs, CT, USA

Kateryna V. Keefer Department of Psychology, University of Western Ontario, London, ON, Canada

Lisa E. Kim School of Psychology, The University of Sydney, Sydney, NSW, Australia

Jeff J. Kosovich Department of Educational Leadership, Policy, and Foundations, University of Virginia, Charlottesville, VA, USA

Rory A. Lazowski K-12 Research, The College Board, Yardley, PA, USA

Anastasiya A. Lipnevich Educational Psychology, Queens College and the Graduate Center, The City University of New York, New York, NY, USA

Carolyn MacCann School of Psychology, The University of Sydney, Sydney, NSW, Australia

Jens Möller Department of Psychology, University of Kiel, Kiel, Germany

Scott Parrigon Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA

Arthur E. Poropat School of Applied Psychology, Griffith University, Mount Gravatt, QLD, Australia

Franzis Preckel Department of Psychology, University of Trier, Trier, Germany

Sandra Prince-Embury The Resiliency Institute of Allenhurst, LLC., West Allenhurst, NJ, USA

Susan E. Rivers Yale Center for Emotional Intelligence, Yale University, New Haven, CT, USA

Richard D. Roberts Center for Innovative Assessments, Professional Examination Service, New York, NY, USA

Rachel Saef Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA

Donald H. Saklofske Department of Psychology, University of Western Ontario, London, ON, Canada

Steven E. Stemler Department of Psychology, Wesleyan University, Middletown, CT, USA

Catalina Torrente Mathematica Public Policy, Yale Center for Emotional Intelligence, Yale University, New Haven, CT, USA

Ulrich Trautwein Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Tübingen, Germany

Kate E. Walton Department of Psychology, St. John's University, Jamaica, NY, USA

Marie C. White School of Education, Nyack College NYC Campus, New York, NY, USA

Sang Eun Woo Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA

Matthias Ziegler Psychological Institute, Humboldt-Universität zu Berlin, Berlin, Germany

Part I
General Background: Theory and Guiding Principles

Chapter 1
Psychosocial Skills: Essential Components of Development and Achievement in K-12

Jeremy Burrus and Meghan Brenneman

1.1 Introduction

As children grow up, they are often told by their parents and teachers that each of them is unique, special, or even "one of a kind." And it's true. Some of us are tall and some of us are short. Some are Elvis people, whereas others prefer the Beatles. Some of us are morning larks and others are night owls (e.g., Roberts & Kyllonen, 1999). In recognition of such differences, for over 100 years, psychologists have attempted to identify and measure the factors that differentiate people and to examine the consequences of these factors for success in education, work, and life.

One individual difference that has received perhaps the most attention by psychologists is cognitive ability (e.g., Carroll, 1993; Horn & Cattell, 1967). A predominant model of cognitive ability states that cognitive ability is comprised of fluid intelligence (e.g., solving problems in novel situations), crystallized intelligence (e.g., knowledge gained by formal education and experience; Horn & Cattell, 1966), and a slew of abilities tied to perceptual processes. Clearly, cognitive ability is important and predicts a number of outcomes. In fact, it is probably the single best predictor of both academic achievement throughout the school years (e.g., Brody, 1997) and job performance (e.g., Schmidt & Hunter, 2004).

"Always remember that you are absolutely unique. Just like everyone else." – Margaret Meade

J. Burrus (✉)
Center for Innovative Assessments, Professional Examination Service,
475 Riverside Drive, Suite 600, New York, NY 10115, USA
e-mail: JBurrus@proexam.org

M. Brenneman
Research and Development, Educational Testing Service,
660 Rosedale Road, Princeton, NJ, USA
e-mail: mbrenneman@ets.org

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_1

There is growing consensus, however, in the realm of public policy and in research in education, psychology, and economics that a number of factors outside of cognitive ability may be just as, or nearly as, important for educational and work success. A few examples include *work ethic*, *creativity*, *teamwork*, *leadership*, and *emotional intelligence*. Because most of these factors often demonstrate low zero-order correlations with cognitive ability, they are often referred to as *noncognitive* factors (see Kyllonen, Lipnevich, Burrus, & Roberts, 2014, for a review). Several alternative terms exist for the factors encompassed by noncognitive factors, including *personal skills*, *personal qualities*, *character traits*, and *psychosocial skills* (Kyllonen et al., 2014). Of these terms, psychosocial skills have seemed to gain the most traction in discussions in the K-12 research arena. Thus, we use this term throughout the remainder of the chapter.

As an organizational framework for our discussion of psychosocial skills, we borrow the prominent organizational taxonomy of personality from the field of personality psychology, the *five-factor model* (also known as the *Big Five*). Personality is "the system of enduring, inner characteristics of individuals that contributes to consistency in their thoughts, feelings, and behavior" (Leary, 2005, p. 3). In essence, personality is a set of characteristics that allow us roughly predict how people's behavior will compare in a range of situations. For example, we might expect an extraverted person to be more talkative than an introverted person at both a party and a library. Furthermore, we may expect a very emotionally stable person to feel less anxiety than a less emotionally stable person during an important school exam and also during their wedding.

For several decades, the field of personality psychology existed without a generally agreed-upon organizational framework, resulting in a multitude of suggested personality traits and a relative lack of coherence in the field. This all changed with the development of the Big Five. The Big Five emerged from factor analyses of people's ratings on the extent to which traits listed in the dictionary were descriptive of them (Tupes & Christal, 1961). The idea of using dictionary trait terms follows from the "lexical hypothesis" or the idea that all of the most important personality characteristics of people should be encoded into their language (e.g., Kyllonen et al., 2014; see also Walton & Billera, Chap. 4, this volume). If a trait is important, then it will be encoded. A strong statement to describe the framework would be that, at least in theory, *every* trait that is important in distinguishing people should be encoded, and thus the Big Five includes *everything* that is important in describing individuals. Importantly, and consistent with the previous statement, the Big Five has been shown to consistently emerge among all age groups (Digman, 1997) and in most countries in the world (see Roberts, Martin, & Olaru, 2015). As such, it works as a suitable organizing framework when describing individual difference.

These Big Five factors of personality are (1) *extraversion* (tendency to be outgoing, assertive, and energetic), (2) *agreeableness* (tendency to be kind, cooperative, and generous), (3) *conscientiousness* (tendency to be organized, responsible, and hardworking), (4) *emotional stability* (tendency to be free from anxiety, worry, and tension), and (5) *openness to experience* (tendency to be imaginative, curious, and insightful). Each of the psychosocial skills we discuss in this chapter can be roughly

categorized into one of these five traits. We acknowledge that several skills are likely blends of several traits and we attempt to identify these blends where possible (see Table 1.1). We place each skill under the trait to which we perceive it to be most highly related. We further divide the Big Five into the three categories of skills we feel are important for students in K-12: *performance skills* ("getting along with school"), *interpersonal skills* ("getting along with others"), and *self-management skills* ("getting along with yourself") (see Kyllonen et al., 2014, for a similar organizational structure). Once again, we acknowledge that the distinctions between these classifications are not always clear-cut. Our intent is to make these distinctions for comprehensiveness and organizational clarity. We conclude with a brief discussion of cross-cultural competence and the malleability of psychosocial skills.

1.2 Performance Skills: "Getting Along" with School

We feel confident in stating that if school were a person, it would want students to perform well. Performance might mean turning in high-quality homework on time and getting good grades on exams. Thus, students who do these things are in essence "getting along with school," as the student's performance is aligned with the school's goals. We place two of the Big Five factors under performance skills: conscientiousness and openness to experience. Psychosocial skills related to these factors are briefly discussed under each.

1.2.1 Conscientiousness and Related Skills

Of the Big Five personality traits, conscientiousness is most consistently related to performance in school (Poropat, 2009). One recent meta-analysis summarized data on the relation of the Big Five personality factors to academic performance from 63 published articles and 17 unpublished dissertations (Poropat, 2009). Results revealed that conscientiousness was a strong predictor of academic achievement (grades) at the primary (e.g., elementary school), secondary (e.g., high school), and tertiary (e.g., college) levels of education, with sample-weighted correlations corrected for scale unreliability of 0.28, 0.21, and 0.23, respectively. Interestingly, at the tertiary level, the conscientiousness/academic achievement and cognitive ability/academic achievement relationships were of equivalent magnitude.

Why should conscientiousness predict academic achievement? The answer to this question is intuitively clear when one considers the lower-order dimensions, or "facets," of which higher-order conscientiousness consists. Several facet-level measures of conscientiousness have been developed, and for the most part, they are consistent with each other (e.g., MacCann, Duckworth, & Roberts, 2009; Peabody & de Raad, 2002; Perugini & Gallucci, 1997; Roberts, Bogg, Walton, Chernyshenko, & Stark, 2004; Roberts, Chernyshenko, Stark, & Goldberg, 2005; Saucier &

Table 1.1 Hypothesized relations of specific psychosocial skills to Big Five personality dimensions

	Big Five dimension				
	Conscientiousness	Openness to experience	Extraversion	Agreeableness	Emotional stability
Performance skills					
Time management	+[a]				
Grit	+[b]		+[b]	+[b]	+[b]
Goal setting	+[c]	+[c]			
Creativity		+[d]			
Curiosity	+[e]	+[e]			
Interpersonal skills					
Leadership	+[f]	+[f]			
Teamwork/collaboration				+[g]	+[g]
Self-management skills					
Self-efficacy					+[f]
Text anxiety					–[h]
Self-esteem	+[i]	+[i]	+[i]	+[i]	+[i]
Coping		+[j]		+[j]	+[j]

[a]Burrus et al. (2013)
[b]Duckworth and Quinn (2009)
[c]Klein and Lee (2006)
[d]McCrae (1987)
[e]Komarraju, Karau, and Schmeck (2009)
[f]Judge, Bono, Ilies, and Gerhardt (2002)
[g]Jensen-Campbell and Graziano (2001)
[h]Moutafi, Furnham, and Tsaousis (2006)
[i]Robins et al. (2001)
[j]MacCann, Lipnevich, Burrus, and Roberts (2012)

Ostendorf, 1999; see also Chap. 7). One particularly comprehensive measure (MacCann et al., 2009) identified eight facets of conscientiousness: (1) *industriousness* (working hard), (2) *perfectionism* (continuing until things are perfect), (3) *tidiness* (being organized), (4) *procrastination refrainment* (getting to work right away), (5) *control* (being self-controlled), (6) *caution* (avoiding taking unnecessary risks), (7) *task planning* (following a schedule), and (8) *perseverance* (continuing to work in the face of pressure or setbacks). The descriptions of these facets should make it clear that several of them ought to be related to academic performance. For instance, students who work hard, get to work right away, and persevere should perform better than those who do not. In a study of 13–19-year-old high school students, MacCann et al. (2009) confirmed that several of the facets of conscientiousness indeed predicted GPA. Specifically, industriousness ($r=0.20$), tidiness ($r=0.14$), control ($r=0.23$), cautiousness ($r=0.20$), task planning ($r=0.13$), and perseverance ($r=0.12$) were all significantly related to GPA. Several of these facets also predicted additional outcomes that might be considered important for students, including class absences and disciplinary infractions.

Time Management Note that one facet of conscientiousness, task planning, could easily be expanded into another psychosocial skill, time management. In a world of constant distractions, the ability to accurately and efficiently prioritize and plan one's activities should be beneficial for the educational success of K-12 students. Time management can be thought of consisting of six important behaviors: (1) having a workspace (being organized and keeping a workspace), (2) meeting deadlines (the extent to which people perceive themselves to be in control of time and to use their time wisely), (3) organizing time and tasks (actions, strategies, and preferred ways of behaving that are associated with successful time management practices, (4) planning ahead (a preference for structure and routine over flexibility, unpredictability, and lack of constraint), (5) setting goals (a sense of purpose, level of focus, and goal-setting capacity), and (6) staying focused (a potential to cope with change and their ability to adapt when change occurs; e.g., Roberts, Krause, & Suk-Lee, 2001). In a study of seventh grade students, Liu, Rijmen, MacCann, and Roberts (2009) found that these six dimensions could be reduced to two dimensions: planning and organization. Each of these dimensions was related to student GPA (planning $r=0.21$; organization $\underline{r}=0.38$). Interestingly, these relationships were just as strong 6 months later (planning $r=0.25$; organization $r=0.42$). Furthermore, a more recent study found that several time management dimensions were significantly related to grades in high school students. Specifically, setting goals ($r=0.31$), planning ahead ($r=0.19$), organizing time and tasks ($r=0.19$), and meeting deadlines ($r=0.24$) were all correlated with grades in ninth grade students (Burrus, Jackson, Holtzman, Roberts, & Mandigo, 2013). Burrus et al. (2013) also showed that five of the six time management facets, with the exception of staying focused ($r=-0.11$), were all significantly correlated with an overall measure of conscientiousness (mean $r=0.50$; range $=0.28$–0.70). We predict that these relationships will only get stronger in the future as technology continues to provide students with an increasing amount of easily accessible, and potentially time-wasting, distractions.

Grit The psychosocial skill that has perhaps received the most recent attention among educators and the popular media is a skill termed grit by Angela Duckworth and her colleagues (e.g., Duckworth, Peterson, Matthews, & Kelly, 2007). In fact, grit was recently featured prominently in both *New York Times Magazine* and a *New York Times* best-selling book (Tough, 2011, 2012). Grit is defined as "perseverance and passion for long-term goals" (Duckworth et al., 2007, p. 1087). The grit scale is intended to measure one's ability to retain interest in goals they set (example item: "My interests change from year to year" (reverse-coded item)) and to persevere in achieving them even in the face of challenges (example item: "Setbacks don't discourage me"). Grit has demonstrated significant relationships with GPA in Ivy League college students ($r = 0.25$) and completion of training courses in West Point cadets ($\beta = 0.48$; Duckworth et al., 2007). Given its strong relationship with conscientiousness ($r = 0.73$–0.77; Duckworth & Quinn, 2009), we expect that, like conscientiousness, grit should also predict academic performance in K-12 students. One way that grit may influence academic performance is through time studying and practicing. Students with more grit simply work harder at mastering material than those with less grit. Consistent with this hypothesis, grit was significantly related to the amount of time national spelling bee contestants (mean age = 13) practiced (Duckworth, Kirby, Tsukayama, Berstein, & Ericsson, 2010).

Goal Setting Most of us can probably recall encountering very different types of students when we were in school. Some students knew what kind of job they wanted when they "grew up," what kind of classes they had to take, and the grades they needed to achieve. They set goals at the beginning of their journey toward these ends because they knew goal setting is one effective way to get where one wants to go. These types of students can be considered self-regulated learners (e.g., Zimmerman, 1990, 2008). Goal setting is just one important aspect of self-regulated learning. Some of the other important components (e.g., time management, self-efficacy) of self-regulated learning are discussed in other sections of this chapter.

A wealth of research evidence suggests that the skills and abilities of self-regulated learning are related to academic achievement at all levels of education (see Zimmerman, 1990, 2008 for reviews). Self-regulated learning refers to "the self-directive processes and self-beliefs that enable learners to transform their mental abilities, such as verbal aptitude, into an academic performance skill, such as writing" (Zimmerman, 2008, p. 166; see Bembenutty, White, & DiBenedetto, Chap. 9, this volume). Zimmerman (2008) has proposed a cyclical model of self-regulation, such that self-regulation occurs before (*forethought phase*), during (*performance phase*), and after (*self-reflection phase*) the learning process. A central process of the forethought phase is setting goals. Goals are "future valued outcomes, the setting of goals is first and foremost a discrepancy creating process. It implies discontent with one's present condition and the desire to attain an object or outcome" (Locke & Latham, 2006, p. 265).

Most educators and leaders of organizations believe that goal setting can be important in helping one perform at a high level, and this belief has been verified by numerous studies in the fields of educational and industrial-organizational psychology

(e.g., Locke & Latham, 1990, 2002, 2006; Schunk, 1990; Zimmerman, Bandura, & Martinez-Pons, 1992). The literature on self-regulated learning and goal setting suggests that the majority of research on goal setting has focused on goal orientation and content. For instance, research on goal orientation has looked at the effect of creating performance (outperforming others) vs. learning (mastery of material) goals (Dweck, 1986). Students with learning goals tend to outperform students with performance goals by using more complex study techniques (Archer, 1994). Research also finds that students with learning goals show better memory for study materials (Graham & Golan, 1991; Pintrich & De Groot, 1990). Furthermore, other research have shown that goal content matters. For example, goals that are difficult to attain tend to lead to better performance than easily attained goals (e.g., Locke & Latham, 1990).

Some aspects of goal setting are related to conscientiousness (e.g., Barrick, Mount, & Strauss, 1993; Klein & Lee, 2006). For instance, in a study of college students, conscientiousness was significantly related both to commitment to one's goals ($r=0.40$) and to being oriented toward learning goals ($r=0.26$; Klein & Lee, 2006). Openness to experience was also significantly related toward being oriented toward learning goals ($r=0.36$), likely because of the intellectual curiosity component of openness to experience (see below).

1.2.2 Openness to Experience and Related Skills

As stated above, openness to experience relates to the tendency to be imaginative, curious, and insightful (see Chap. 5). Facets of openness to experience include *fantasy* (include having a vivid imagination), *aesthetics* (appreciation of the arts), *feelings* (intensity of feeling), *actions* (preference for variety), *ideas* (enjoying complex problems over simple ones), and *values* (questioning of conventional norms; Costa & McCrae, 1992). Although the relationships are not as strong as with conscientiousness, the Poropat (2009) meta-analysis found that openness to experience was related to grades at the primary ($\rho=0.24$), secondary ($\rho=0.12$), and tertiary levels of education ($\rho=0.07$). The relationship of openness to experience to academic achievement may be due to the fact that openness tends to be moderately related to cognitive ability, and thus highly open individuals tend to get good grades via higher intelligence (Ackerman & Heggestad, 1997). As such, openness to experience and its related psychosocial skills are the most highly "cognitive" of the noncognitive psychosocial skills we discuss in this chapter. Two psychosocial skills related to openness to experience include *creativity* and *curiosity*.

Creativity Over the past decades of research on creativity in psychology, a standard, universally agreed-upon definition of creativity has remained elusive (see Kaufman, Beghetto, & Dilley, Chap. 6, this volume). Typically, when people talk about creativity, they mean the ability to develop original and useful ideas (Runco & Jaeger, 2012). It is typically thought to involve two processes. Divergent thinking involves the ability to think of many possible solutions to a problem, whereas convergent thinking

involves the process of combining new and old information to come up with a new solution. Because creativity involves thinking, people often assume that it is highly related to cognitive ability. However, a meta-analysis examining the relationships of many types of tests of creativity to IQ scores revealed only a mean correlation of 0.17 (Kim, 2005). In fact, creativity seems to be just as, if not more highly, related to openness to experience as to cognitive ability (e.g., McCrae, 1987; Feist, 1998). For example, one study found that the facets of openness to experience correlated with tests of divergent thinking at $r = 0.17–0.31$. Furthermore, a meta-analysis revealed that very creative scientists and artists were more open to experience than less creative scientists and artists (Feist, 1998).

Curiosity We suspect that most teachers want their students to be curious because curious students should be motivated to learn on their own outside of class simply for personal gratification. Although there are several models of curiosity, all agree that curiosity "…is an approach-oriented motivational state associated with exploration" (Kashdan & Silvia, 2009, p. 368). People who are curious seek out information in order to learn it. Kashdan and Silvia define curiosity as "…the recognition, pursuit, and intense desire to explore novel, challenging, and uncertain events" (2009, p. 368). Perhaps not surprisingly, a meta-analysis of studies employing students from fifth through twelfth grade has found that constructs related to curiosity (e.g., liking for subject areas, desire for more education) are a good predictor of grades, with sample-weighted correlations ranging from 0.17 for literature grades to 0.35 for science grades. In the psychological and educational psychology literatures, several constructs are studied that could be considered to be conceptually equivalent, or at least strongly related to, curiosity. These include intrinsic motivation, need for cognition, and typical intellectual engagement (TIE; see Jebb, Saef, Parrigon, & Woo, Chap. 5, this volume).

Intrinsic motivation can be thought of as "the desire to engage in behaviors for no reason other than sheer enjoyment, challenge, pleasure, or interest," whereas the desire to engage in behaviors to impress others, earn grades or money, etc. is referred to as *extrinsic motivation* (Lepper, Corpus, & Iyengar, 2005, p. 184). A recent study found both types of motivation to be significantly related to conscientiousness, but only intrinsic motivation to be significantly related to openness to experience (Komarraju et al., 2009). Several studies have shown that intrinsic motivation is positively related to academic achievement (e.g., Gottfried, 1985; Lepper et al., 2005; Lloyd & Barenblatt, 1984; see Hulleman, Barron, Kosovich, & Lazowski, Chap. 10, this volume). In one study, researchers measured the intrinsic (example item: "I work on problems to learn how to solve them") and extrinsic (example item: "I work on problems because I'm supposed to") motivation of nearly 800 third through eighth grade students (Lepper et al., 2005). Results revealed that intrinsic motivation was positively and significantly correlated with grades ($r = 0.34$), and extrinsic motivation was negatively and significantly correlated with grades ($r = -0.23$). Intrinsic motivation was also related to age, as students at each grade level demonstrated lower intrinsic motivation than students in each of the earlier grades. Extrinsic motivation, by contrast, was lower for fourth grade students than for third grade students but there were relatively no grade differences from fourth to

eighth grade. One possibility for the finding of seemingly diminishing intrinsic motivation of students as they get older is the *overjustification effect* (e.g., Deci, Koestner, & Ryan, 1999) or the finding that the presence of external rewards can serve to diminish intrinsic motivation. As such, the desire to please adults by getting good grades may reduce students' motivation to learn for their own edification as they get older.

Two additional constructs are worth mentioning in the context of curiosity. *Need for cognition* (NFC) is the tendency to prefer effortful mental activities (Cacioppo & Petty, 1982). That is, students who are high in the need for cognition enjoy thinking through difficult problems (see Jebb, Saef, Parrigon, & Woo, Chap. 5, this volume). NFC has been demonstrated to be strongly related to openness to experience (McCrae, 1996) and also predicts academic performance (Preckel, Holling, & Vock, 2006; Sadowski & Gulgoz, 1992). Typical intellectual engagement (TIE; Goff & Ackerman, 1992) is a construct that is nearly indistinguishable from NFC (von Stumm, Hell, & Chamorro-Premuzic, 2011; Woo, Harms, & Kuncel, 2007) that refers to people's preference to engage in mental activities that are difficult and taxing. As with NFC, it also demonstrates a strong relation with openness to experience (McCrae, 1996) and strongly predicts academic achievement (e.g., von Stumm et al., 2011). Taken together, the evidence suggests that curiosity and related constructs are an important component of student success. Importantly, meta-analysis has demonstrated that TIE predicts academic performance even after controlling for intelligence and conscientiousness (von Stumm et al., 2011).

1.3 Interpersonal Skills: Getting Along with Others

The way that work gets done in schools and at work is changing. Teamwork and collaboration skills are becoming increasingly important as group work becomes more emphasized in today's schools and organizations. This new emphasis is reflected in several of the "21st century skills" frameworks that have been developed with the goal of outlining the skills that students should have to be effective workers and citizens in the twenty-first-century workforce. For example, the *Partnership for 21st Century Skills* (P21) framework contains several skills that involve getting along with others or society as a whole (Partnership for 21st Century Skills, 2012). These include *leadership* and *teamwork/collaboration*. Below we discuss these issues in terms of their relationships to the Big Five components extraversion and agreeableness.

1.3.1 Extraversion and Related Skills

Extraversion is particularly related to sociability, as it involves the tendency to be outgoing, assertive, and energetic. Facets of extraversion include *warmth*, *gregariousness*, *assertiveness*, *social ascendancy*, and *forcefulness of expression* (Costa & McCrae, 1992). In terms of academic performance, extraversion shows small

relations with performance in primary students ($\rho = 0.18$) but is unrelated to performance in secondary ($\rho = -0.03$) and tertiary ($\rho = -0.01$) students (Poropat, 2009). Although extraversion does not seem to relate to performance in the classroom, it still may have important implications for student behaviors both in and out of the classroom. Some of these behaviors can be predicted by inspecting extraversion's facets. For instance, the gregariousness, assertiveness, and forcefulness of expression facets suggest that extraverts should participate in class discussion to a greater extent than introverts, a prediction that has been verified in research on college students (Caspi, Chajut, Saporta, & Beyth-Marom, 2006). This suggests that the relationships in the meta-analysis of Poropat (2009) may be moderated by class grading criteria, such that extraverts may perform better in classes that incorporate participation grades than in classes that do not.

Being highly extraverted has other upsides. One might also predict from the warmth and gregariousness facets that extraverts should have more friends and larger social networks in general. This prediction has also been borne out by research (Asendorpf & Wilpers, 1998; Selfhout et al., 2010). One longitudinal study on college students found that extraversion predicted the number of peer relationships students claimed to have ($r = 0.34$) and it continued to predict number of peer relationships 12 months later ($\beta = 0.19$). Furthermore, sociability and shyness predicted whether the student was in love 12 months later ($\beta = 0.16$ and -0.20, respectively; Asendorpf & Wilpers, 1998). Extraverts also tend to have friends who are also similarly extraverted (Selfhout et al., 2010). Because quantity and quality of social relationships strongly predict happiness (Diener & Seligman, 2002), we might also expect extraverts to be happier on average than introverts. Meta-analysis finds that they are ($r = 0.28$, Steel, Schmidt, & Shultz, 2008).

Every upside has its downside. Although the social aspects of extraversion may ultimately lead to more friends and happiness, it may also lead to negative health outcomes. Specifically, extraverts tend to engage in riskier behaviors (e.g., Nicholson, Soane, Fenton-O'Creevy, & Willman, 2005). For example, research has found that extraverted university students drink more alcohol and drink and drive more than introverted students (Vollrath & Torgersen, 2002). Much evidence suggests that extraverted youth and adolescents drink more alcohol in order to increase positive affect (Kuntsche, Knibbe, Gmel, & Engels, 2006). In essence, extraverts like excitement and they drink to feel that way.

Leadership Leadership is often mentioned as a skill that is important for students to develop. Although people tend to think they know what leadership is, it is often difficult to define, much less measure. One view is that leaders can be either transactional or transformational (e.g., Bass, 1998). On the one hand, leaders with a transactional style tend to lead others by exchanging rewards (e.g., money, good grades) for effort, and they tend to intervene in a group's activities only when standards for performance are not being met. On the other hand, leaders with a transformational style tend to lead others by providing a vision for others and a mission to try to accomplish. They lead by inspiring people and pay a lot of close attention and mentoring to team members. Not surprisingly, transformational

leadership is usually considered more effective and more desirable of the two leadership styles, and some evidence suggests that adolescents can learn to be transformational leaders by interacting with adults who exhibit these behaviors (Zacharatos, Barling, & Kelloway, 2000). Specifically, one study found that both female and male adolescents whose fathers used a transformational leadership style subsequently used a transformational leadership style themselves when interacting with their peers (Zacharatos et al., 2000). Meta-analysis has found that extraversion is the strongest personality predictor of leadership (as indexed by emergence and effectiveness) in adults, while conscientiousness and openness are also consistent predictors (Judge et al., 2002). Note, however, that this study does not speak to personality predictors of leadership style, although one study did find a significant relationship between extraversion and using an inspirational persuasive style (Cable & Judge, 2003).

1.3.2 Agreeableness and Related Skills

When someone says a student is a "nice kid," they typically mean that he or she is agreeable. Facets of agreeableness include *trust*, *morality*, *altruism*, *cooperation*, *modesty*, and *sympathy* (Costa & McCrae, 1992). Similar to extraversion, agreeableness demonstrates a relation to academic performance at the primary level of schooling ($\rho = 0.30$) but smaller relations at the secondary ($\rho = 0.05$) and tertiary ($\rho = 0.06$) levels (Poropat, 2009). Also similar to extraversion, though, a person's level of agreeableness may have important implications for behavior both in and out of the classroom. For example, in one study, agreeableness in fourth grade boys was associated with rates of juvenile delinquency (e.g., shoplifting, vandalism, breaking and entering, selling drugs), with nondelinquent students higher on agreeableness than delinquent ones (John, Caspi, Robins, Moffitt, & Stouthamer-Loeber, 1994). Consistent with these findings, another study found that agreeableness was significantly correlated with conduct problems (e.g., picking fights, suspension from school) in children age 9–13 ($r = -0.78$; Ehrler, Evans, & McGhee, 1999). These issues seem to persist into adulthood, as further research has revealed that people who are low in agreeableness at age 8 who remained low in agreeableness at age 33 experienced higher rates of alcoholism, arrests, depression, and less career stability than their study peers (Laursen, Pulkkinen, & Adams, 2002). The findings that demonstrate that agreeableness is associated with conduct problems may suggest that agreeableness should thus be associated with *teamwork/collaboration*. Clearly, getting into fights is not an example of good teamwork behavior.

Teamwork/Collaboration If five people were asked if teamwork is an important skill for students to develop, odds are likely that all five would say "absolutely." However, if you asked each to define the essential elements of teamwork, odds are likely that you would get five different answers. Within the vast scope of educational research on psychosocial skills, very few have tried to define and measure teamwork.

One group of researchers who have attempted to do so stated that teamwork consists of four essential components: (1) cooperation with other team members, (2) influencing team members through support and encouragement, (3) resolution of conflicts among team members via negotiation strategies, and (4) guidance and mentorship of other team members (Wang, MacCann, Zhuang, Liu, & Roberts, 2009). Wang et al. developed a measure of teamwork for high school students, with example items including "I enjoy bringing team members together (cooperation item)," "I like to be in charge of groups or projects (guidance and mentorship item)," and "I take others' interests into account (resolution of conflicts item)." Responses on the teamwork scales were correlated with student performance on several high school classes. Interestingly, responses were most strongly and consistently related to the class that most required the most teamwork, music class. Specifically, the cooperation, guidance and mentorship, and resolution of conflicts scales all significantly predicted music grades, with $rs = 0.38$, 0.40, and 0.50, respectively.

Beyond the work of Wang et al. (2009), little research has attempted to assess and study teamwork as a whole. There has, however, been some research conducted on children on constructs that should be critical to teamwork. For example, in a sample of sixth and eighth grade students, agreeableness was related to endorsements of positive conflict resolution styles (e.g., compromise) and rejection of negative resolution styles (e.g., use physical force, Jensen-Campbell & Graziano, 2001). Of the remaining personality traits, only emotional stability related to endorsements, such that students low in emotional stability tended to endorse negative resolution styles. More recent work on young adults has found that agreeableness predicts prosocial behaviors that should be beneficial to teamwork (e.g., sharing, helping, empathizing; Caprara, Alessandri, & Eisenberg, 2012). In addition, the relationship between agreeableness and prosocial behaviors was mediated by self-transcendence values (values that emphasize equality and concern for others) and the belief that one has the ability to be empathic.

1.4 Self-management Skills: Getting Along with Yourself

In addition to students performing well in school and getting along with others, we want students to possess and develop self-management skills (see also Bembenutty et al., Chap. 9, this volume). Depending on the context, self-management can be conceptualized in many different ways; however, for the purpose of this chapter, we define it as the methods, skills, and strategies individuals apply to direct their behavior toward achieving a desired objective or goal. When placed in the context of school, self- management might mean feeling confident and fighting self-doubt while taking an exam and not allowing daily stressors to negatively impact academic performance. We think exhibiting strong self-management skills requires an introspective knowledge of one's behavior and the ability to not be overcome by one's fears and doubts (i.e., knowing and getting along with oneself). Therefore, we place emotional stability as the one Big Five factors under self-management skills and first

briefly describe this factor. We then identify the following psychosocial skills related to emotional stability: self-efficacy, test anxiety, self-esteem, and coping.

1.4.1 Emotional Stability

When you google the term "emotional stability," the results are extensive and immediate. The dictionary defines emotional stability as the capacity to maintain one's emotional balance under stressful circumstances. Therefore, a person, who is emotionally stable, as opposed to someone who is neurotic, is able to better tolerate the day-to-day strains and stresses by not getting upset, anxious, or angry than a person who is less emotionally stable. The literature suggests relationships between emotional stability, self-efficacy, subjective well-being, and academic performance. In the meta-analytic study cited throughout this chapter, emotional stability shows small relations with performance in primary students (ρ=0.20) but is unrelated to performance in secondary (ρ=0.01) and tertiary (ρ=−0.01) students (Poropat, 2009).

Emotional stability is associated with self-efficacy (Judge et al., 2002), which in turn is positively correlated with academic performance (see the following section). Emotional stability is also found to be related to happiness and life satisfaction with correlations of r=−0.46 and r=−0.38, respectively (Steel et al., 2008).

Self-efficacy If a student believes they are good at math, will they perform better? We all have come across individuals we would categorize as exhibiting confidence in their academic abilities, but does this matter when we think about performance in school? Variations in performance can now be attributed, in part, to differences among people in their beliefs and perceptions about their ability (Lennon, 2010; see Bembenutty et al., Chap.9, this volume). The concept of self-efficacy or a person's belief in his or her ability to succeed in specific situations (Bandura, 1977) is now being incorporated into studies examining children's social and academic achievements. On the one hand, people with high self-efficacy approach difficult tasks with fervor, feel confident while doing so, and exhibit a stick-to-itiveness. On the other hand, people with low self-efficacy may believe that things are tougher than they really are, a belief that fosters stress, and can restrict problem solving (Pajares, 1996).

The relationship between self-efficacy and academic performance has been widely investigated. In their meta-analysis, Multon, Brown, and Lent (1991) summarized 36 studies looking at the relationship between self-efficacy and academic performance and found an overall effect size of r=0.38. On an international scale, one of the measurement initiatives using self-efficacy was conducted by the Organisation for Economic Co-operation and Development (OECD) in a study of 4000 15-year-olds from 25 countries. Perceived self-efficacy was one of the 14 factors examined, and scores indicated perceived self-efficacy demonstrated modest associates with standardized reading and math performance (r=0.28 and 0.29, respectively; Marsh, Hau, Artelt, & Peschar, 2006).

Self-efficacy can also help predict college-level outcomes. Robbins et al. (2004) conducted an extensive meta-analysis (109 studies) on the relationship between

psychosocial and study skill factors (PSFs) and college outcomes. Included in their list was academic self-efficacy, which proved to be one of the strongest predictors of persistence of all PSFs examined, and moderately predicted GPA (mean true-score correlation 0.50). Furthermore, Richardson, Abraham, and Bond's (2012) meta-analysis found that academic self-efficacy ($\rho = 0.28$) and performance self-efficacy ($\rho = 0.67$) were both significant predictors of college GPA. The large relationship of performance self-efficacy and GPA suggests that performance self-efficacy may simply act as a proxy for achievement (e.g., I believe I am good at math because I get good grades in math).

Text Anxiety It makes sense that highly neurotic individuals are more likely to experience greater levels of anxiety than emotionally stable people. In school, one of the greatest sources of anxiety manifests during tests. Not surprisingly, research has found that test anxiety is correlated with emotional stability (e.g., Moutafi et al., 2006). Meta-analyses have consistently found negative relationships between test anxiety and academic performance. For example, Ackerman and Heggestad (1997) reported a negative correlation between test anxiety and learning and memory, and Seipp (1991) reported a negative correlation between anxiety and academic performance. Research also has shown that highly test-anxious students are more likely to engage in off-task behaviors while taking a test such as looking around the room or looking at the teacher (Nottelman & Hill, 1977; Prins, Groot, & Hanewald, 1994). In another meta-analysis, Hembree (1988) identified 73 studies that compared high- and low-test-anxious students on academic measures (e.g., IQ, achievement, and aptitude tests) and found an effect size of −0.48 for academic achievement, signifying that students high in test anxiety had lower achievement than students low in test anxiety. These studies all use correlational data to identify the relationship between anxiety and achievement or other factors. It should be noted that while causation cannot be implied by correlational studies, this research suggests that a relationship may exist. However, some causal studies in test anxiety have been completed. For example, Hembree (1988) also found that in programs that reduced test anxiety, students' scores improved, showing that reduction of test anxiety positively impacted test performance.

Finally, research has suggested that students high in test anxiety perform poorly in evaluative situations in part because they focus too much attention on task-irrelevant stimuli. In support of this contention, Dusek, Mergler, and Kermis (1976) found that highly test-anxious students improved in test performance on a memory task when they were provided with instructions that focused their attention on the task (e.g., asking them to overtly label pictures), while students that were not highly test anxious did not improve as a function of the instructions (for an overview on academic emotions and their regulation, see Goetz & Biek, Chap. 11, this volume).

Self-esteem Related to self-efficacy and self-concept but unique in its own right, self-esteem is the overall emotional evaluation of one's self-worth or a value judgment of oneself. Individuals who are more emotionally stable and less neurotic would hold a favorable evaluation of the self, while those less emotionally stable and more neurotic would hold an unfavorable definition of the self. Self-esteem has

been shown to correlate positively with all of the Big Five personality traits. In a large study of approximately 325,000 participants, Robins, Tracy, Trzesniewski, Potter, and Gosling (2001) found that self-esteem correlated positively with extraversion ($r=0.38$), agreeableness ($r=0.13$), conscientiousness ($r=0.24$), emotional stability ($r=0.50$), and openness ($r=0.17$). Baumeister, Campbell, Krueger, and Vohs (2003) explain that self-esteem has been an important factor to understand in schools for the plausible thinking that high self-esteem will lead to good schoolwork, setting of high aspirations, persisting through failure, and confidence in problem solving. The self-esteem movement is so popular in schools it found a place in many character education programs. The state of California even created the Task Force to Promote Self-Esteem and Personal and Social Responsibility, whose guiding principle was to raise young people's self-esteem in order to increase the number of socially responsible people in society, though this movement received a lot of criticism in the press. Empirically, the evidence suggesting a relationship between self-esteem and academic performance is mixed. Many correlational studies have indicated a positive correlation between self-esteem and various measures of academic performance (Bowles, 1999; Davies & Brember, 1999; Hansford & Hattie, 1982; Kugle, Clements, & Powell, 1983; Wylie, 1979) yet others have drawn different conclusions (Zimmerman, Copeland, Shope, & Dielman, 1997). Baumeister et al. (2003) conclude that the positive but weak correlations do not differentiate whether self-esteem is a cause or a result of school performance. Causal studies investigating self-esteem and academic performance have yielded different results as most of the evidence suggests that self-esteem has no impact on subsequent academic achievement (Bachman & O'Malley, 1977, 1986; Maruyama, Rubin, & Kingsbury 1981) but any relationship is due to an unobserved or unknown third variable (Pottebaum, Keith, & Ehly, 1986).

So why are schools consumed with building students' self-esteem despite a weak or insignificant relationship between self-esteem and academic performance? Perhaps the answer can be found in the link between self-esteem and happiness. Baumeister et al. (2003) stated people with high self-esteem are significantly happier than other people. In their 1998 meta-analysis of 137 personality traits and happiness (labeled subjective well-being), DeNeve and Cooper (1998) found that private collective self-esteem was one of the strongest predictors of happiness ($\rho=0.31$). Additionally, Diener and Diener (1995) found a correlation between self-esteem and happiness ($r=0.47$) in their study with 13,000 college students in 31 countries.

Finally self-esteem has also been found to have some relationship to coping (see next section) as high self-esteem has shown to improve persistence in the face of failure, though this often depends on the degree of the stressor (Baumeister et al., 2003).

Coping The everyday life of a modern child can be stressful. The daily barrage of school, work, family, peers, and, more recently, social media creates stress at varying levels of intensity. Understanding how students respond to challenges or stressful situations is important as it can impact students' performance in school. There are several distinctions between coping strategies, yet three approaches that

emerge often in the literature are problem-focused coping, through resolving the root cause of the stressful situation; emotion-focused coping, through a focus on one's emotional responses to the stressor; and avoidant coping, through avoiding the stressor as much as possible (MacCann et al., 2012). Coping strategies have been linked with personality traits, life satisfaction, academic achievement, and a range of well-being measures (e.g., Carver & Connor-Smith, 2010; Diener, Lucas, & Napa Scollon, 2006; MacCann, Fogarty, Zeidner, & Roberts, 2011; MacCann et al., 2012). For example, MacCann et al. (2012) found that all three coping styles were significantly related to emotional stability, agreeableness, and openness to experience.

In their study of 354 high school students, MacCann et al. (2012) investigated the relationship between coping and high school grades, life satisfaction, and students' positive and negative reactions toward school. The authors found that problem-focused coping predicts grades, life satisfaction, and greater positive feelings toward school. Avoidant and emotion-focused coping predicted negative feelings toward school. Interestingly, the authors also found that avoidant coping predicted positive feelings toward school suggesting that avoidant coping can be an effective strategy in some situations.

1.5 What's Missing? Cross-Cultural Competence (3C)

Space limits permit us from discussing all important psychosocial skills. One specific skill that we have yet to discuss that we believe will be increasingly important for students is *cross-cultural competence* (3C). It would seem that the need for 3C will grow as the world becomes increasingly interconnected and globalized. That is, innovations in travel and communications have led to more interactions among people of different cultures than ever before. Along with this increasing interaction also comes a greater chance of misunderstanding stemming from cultural differences. Such misunderstandings can be avoided by culturally competent individuals. There are many alternative terms for 3C (e.g., *intercultural effectiveness*, *intercultural competence*, *cultural intelligence*), but each for the most part states that an individual, when interacting with a person or people from another culture, should have "(a) the ability manage psychological stress, (b) the ability to communicate effectively, and (c) the ability to establish interpersonal relationships" (Chiu, Lonner, Matsumoto, & Ward, 2013, p. 843). This three-part definition suggests that several Big Five personality components may be associated with 3C. Consistent with this assertion, meta-analysis has found that emotional stability, extraversion, agreeableness, and conscientiousness are all significantly associated with expatriate job performance (Mol, Born, Willemsen, & van der Molen, 2005).

A more recent qualitative synthesis of 3C frameworks revealed near consensus that several more specific skills should be related to 3C (Burrus, Brenneman, Carney, Ezzo, & Roberts, 2014). These include self-efficacy, cultural awareness, self-monitoring, tolerance for ambiguity, self-regulation, flexibility, influencing skills, and cognitive complexity (e.g., the ability to see subtleties and nuance; e.g., Bieri, 1955). To sum these up, these skills essentially allow for one to properly attribute different or unexpected

behaviors or communication problems to cultural differences and adjust one's behavior in order to function well in the situation. In addition, it involves the ability to handle the stress inherent when undergoing a cultural conflict, thus avoiding "culture shock."

What skills might be important for students to have for developing 3C? We posit that one important skill we have yet to discuss is *metacognition*, or "thinking about one's knowledge" (Klafehn, Li, & Chiu, 2013). Metacognition should be important in the development of 3C because it is the process which allows one to test their cultural assumptions and adjust them if necessary in order to create a new, and more useful, mental model of the world. For example, a student studying abroad may become disoriented due to experiencing unfamiliar culturally determined behavioral norms. Metacognition should allow this student to compare this new experience with his/her cultural assumptions about human behavior and update his/her knowledge and assumptions, thus becoming more culturally aware and cross-culturally competent. Note that a similar process can occur with nearly all learning (e.g., Efklides, 2006).

1.6 Can Psychosocial Skills Be Improved?

Our use of the five-factor model of personality as our organizing structure for psychosocial skills might give many readers pause because it has been traditionally thought that personality is fixed; and, if personality is fixed, then there may be nothing that can be done to improve psychosocial skills. Fortunately, an accumulation of evidence suggests that personality changes much more than previously thought throughout the lifespan (see Chap. 4). In addition, research shows that many psychosocial skills can be improved through intervention. Below, we touch very briefly on some of these topics.

Longitudinal research shows that personality at the broad domain level changes as people get older. Roberts, Walton, and Viechtbauer (2006) conducted a meta-analysis of 92 studies which demonstrated that people tend to increase in agreeableness, conscientiousness, emotional stability, and social dominance (a trait related to leadership) as they age. If these traits change, it follows that intentionally changing them is at least within the realm of possibility. Of course, it is important to mention that by "changing personality," we mean simply helping people improve the specific skills that will help them succeed in school and work.

In locating the specific skills that will contribute to success in school and work, a closer inspection of personality change at the narrower facet level may be informative. That is, some personality facets might contribute more to success than others. Take the conscientiousness scale of MacCann et al. (2009) described above; whereas industriousness significantly predicted GPA ($r = 0.20$), perfectionism did not ($r = 0.07$). Thus, an intervention targeting industriousness will theoretically be more useful than an intervention targeting perfectionism. In addition, the fact that behaviors at the facet level are more specific and focused in nature may lend themselves to targeted interventions. For example, it should be easier to intervene on students' task planning (e.g., making schedules, following directions) than on their overall level of conscientiousness.

Some cross-sectional research has found that there are indeed age differences in facets of personality (Soto, John, Gosling, & Potter, 2011). Although ideally this research would have a longitudinal design, it does provide some evidence that personality also changes at the facet level. More specifically, in general, older individuals are higher than younger people in the conscientiousness facets of self-discipline and order, the agreeableness facets of altruism and compliance, and the openness facets of ideas and aesthetics. Conversely, older individuals are lower than younger people in the emotional stability facets of anxiety and depression and the extraversion facets of assertiveness (somewhat inconsistent with Roberts et al., 2006) and activity. It is important to note that there is a noticeable dip that occurs between the ages of 10 and 15 such that people become less conscientious, less agreeable, less extraverted, less open, and less emotionally stable. If this dip is found to be associated with problems in school, this may prove to be a crucial period of time in which to intervene upon students.

The research described above shows that personality traits, and hopefully associated psychosocial skills, do indeed change over time. Although this suggests that it is possible to improve psychosocial skills, it does not address the question of whether psychosocial skills are responsive to interventions. There is emerging evidence, however, that programs designed to improve students' psychosocial skills can be effective. For example, social and emotional learning (SEL) programs have been demonstrated to help students improve their ability to recognize and manage emotions, set goals, and work well with others (see Chaps. 11 and 13). A meta-analysis conducted by the Collaborative for Academic, Social, and Emotional Learning found that SEL programs had a positive influence on K-8 grade students' social and emotional skills ($d=0.60$), attitudes toward self and others ($d=0.23$), prosocial behavior ($d=0.24$), and emotional distress ($d=0.23$; Payton et al., 2008). Importantly, students who took part in these programs also improved their academic performance ($d=0.28$).

Additionally, other interventions in schools have demonstrated promise in improving psychosocial skills. Classroom interventions designed to improve critical thinking skills have been shown to be effective in improving metacognitive awareness and monitoring of one's learning of reading passages (Haller, Child, & Walberg, 1988). Furthermore, study skills interventions have been found to improve both student motivation and academic performance (Dignath, Buettner, & Langfeldt, 2008). Finally, interventions designed to reduce the negative emotionality and worry that comes with test anxiety have been demonstrated to improve test performance (Hembree, 1988).

The work of these researchers, and others, demonstrates that it is possible to improve psychosocial skills through intentional interventions. As we noted above, although we only briefly review this literature, and thus omit much of the work on psychosocial interventions, more work is needed examining the efficacy of psychosocial skills interventions.

1.7 Conclusion

We hope that this chapter has demonstrated the importance for students to possess a wide array of psychosocial skills. If we have done so, we will have considered our efforts a success. We do feel, however, that we have left much work to be done. Important work remains in several areas. Some of these are outlined in more detail below:

- *Identifying which skills are the most essential skills for students*. We have discussed 13 psychosocial skills that we feel, and evidence suggests, are important for students to develop. This evidence, however, does not speak to their relative importance. Are there one or two skills that are essential to develop whereas others, while nice to have, are not necessary? Resources, limited as they are, parents, students, and teachers can only focus on fostering a certain number of skills, and knowing which skills are most important would help sharpen this focus.
- *Identifying which psychosocial skills are most amenable to intervention*. Related to the point above, it will also be important to identify which psychosocial skills can be intervened upon. For instance, perhaps it is indeed the case that "leaders are born, not made." If so, then trying to intervene on leadership would be futile. On the other hand, it may be easy to teach students how to manage their time, and thus time management interventions should be considered.
- *Creating a parsimonious taxonomy of psychosocial skills*. In the current chapter, we use the Big Five personality taxonomy as an organizing structure of psychosocial skills. This is not the only possible structure that could be used, however. For example, the National Research Council (Koenig, 2011) divided psychosocial skills into cognitive skills (e.g., critical thinking), interpersonal skills (e.g., social skills), and intrapersonal skills (e.g., time management). Both structures, however, might be too broad to be useful in suggesting new psychosocial skills to consider. As the periodic table of the elements reveals missing elements yet to be discovered, a detailed organizing taxonomy of psychosocial skills might reveal important psychosocial skills yet named.
- *Developing new assessments of these skills that are resistant to faking*. Typically, psychosocial skills are assessed through self-report means. For example, "I enjoy working in teams" might be a typical item measuring teamwork. If a respondent is trying to provide the "right" answer, he or she will easily be able to report that he or she enjoys working in teams even if this is not really the case. This "faking" issue can be a problem for the assessment of psychosocial skills, whether the assessment is used for selection or development (see Ziegler, MacCann, & Roberts, 2011 for an extensive review). Some promising options include the use of forced-choice methods (e.g., Stark, Chernyshenko, Drasgow, & White, 2012) and serious video games (e.g., Shute, Ventura, Kim, & Wang, 2014).

It should be clear from the above points (and from the entire chapter as a whole) that, while psychosocial skills are important for success in school (and work), there is still much work to be done in identifying, classifying, measuring, and training

them. Improvements in each of these areas can, and should, lead to important benefits for students and society as a whole.

Acknowledgment All statements expressed in this article are the author's and do not reflect the official opinions or policies of ProExam or Educational Testing Service. Correspondence concerning this chapter should be directed to the first author.

References

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121*, 219–245.

Archer, J. (1994). Achievement goals as a measure of motivation in university students. *Contemporary Educational Psychology, 19,* 430–446.

Asendorpf, J. B., & Wilpers, S. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology, 74*, 1531–1544.

Bachman, J. G., & O'Malley, P. M. (1977). Self-esteem in young men: A longitudinal analysis of the impact of educational and occupational attainment. *Journal of Personality and Social Psychology, 35*, 365–380.

Bachman, J. G., & O'Malley, P. M. (1986). Self-concepts, self-esteem, and educational experiences: The frog pond revisited (again). *Journal of Personality and Social Psychology, 50*, 35–46.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191–215.

Barrick, M. R., Mount, M. K., & Strauss, J. P. (1993). Conscientiousness and performance of sales representatives: Test of the mediating effects of goal setting. *Journal of Applied Psychology, 78*, 715–722.

Bass, B. M. (1998). *Transformational leadership*. Mahwah, NJ: Erlbaum.

Baumeister, R., Campbell, J., Krueger, J., & Vohs, K. (2003). Does high self-esteem cause better performance, interpersonal success, happiness or healthier lifestyles? *Psychological Science in the Public Interest, 4*, 1–44.

Bieri, J. (1955). Cognitive complexity-simplicity and predictive behavior. *Journal of Abnormal and Social Psychology, 51*, 263–268.

Bowles, T. (1999). Focusing on time orientation to explain adolescent self concept and academic achievement: Part II. Testing a model. *Journal of Applied Health Behaviour, 1*, 1–8.

Brody, N. (1997). Intelligence, schooling, and society. *American Psychologist, 52*, 1046–1050.

Burrus, J., Brenneman, M. W., Carney, L. C., Ezzo, C., & Roberts, R. D. (2014). *Development of an assessment of cross-cultural competence (3C): Expanded literature review.* Unpublished manuscript.

Burrus, J., Jackson, T., Holtzman, S., Roberts, R. D., & Mandigo, T. (2013). *Examining the efficacy of a time management intervention for high school students* (ETS RR-13-25). Princeton, NJ: Educational Testing Service.

Cable, D. M., & Judge, T. A. (2003). Managers' upward influence tactic strategies: The role of manager personality and supervisor leadership style. *Journal of Organizational Behavior, 24*, 197–214.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*, 116–131.

Caprara, G. V., Alessandri, G., & Eisenberg, N. (2012). Prosociality: The contribution of traits, values, and self-efficacy beliefs. *Journal of Personality and Social Psychology, 102*, 1289–1303.

Carroll, J. B. (1993). *Human cognitive abilities*. Cambridge, MA: Cambridge University Press.

Carver, C. S., & Connor-Smith, J. (2010). Personality and coping. *Annual Review of Psychology, 61*, 679–704.

Caspi, A., Chajut, E., Saporta, K., & Beyth-Marom, R. (2006). The influence of personality on social participation in learning environments. *Learning and Individual Differences, 16*, 129–144.

Chiu, C. Y., Lonner, W. J., Matsumoto, D., & Ward, C. (2013). Cross-cultural competence theory, research, and application. *Journal of Cross-Cultural Psychology, 44*, 843–848.

Costa, P. T. J., & McCrae, R. R. (1992). *Revised NEO personality inventory and NEO five-factor inventory: Professional manual*. Lutz, FL: Psychological Assessment Resources.

Davies, J., & Brember, I. (1999). Reading and mathematics attainments and self-esteem in years 2 and 6—an eight-year cross-sectional study. *Educational Studies, 25*, 145–157.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*, 627–668.

DeNeve, K. M., & Cooper, H. (1998). The happy personality: A meta-analysis of 137 personality traits and subjective well-being. *Psychological Bulletin, 124*, 197–229.

Diener, E., & Diener, M. (1995). Cross-cultural correlates of life satisfaction and self-esteem. *Journal of Personality and Social Psychology, 68*, 653–663.

Diener, E., Lucas, R. E., & Napa Scollon, C. (2006). Beyond the hedonic treadmill: Revising the adaptation theory of well-being. *American Psychologist, 61*, 305–314.

Diener, E., & Seligman, M. E. P. (2002). Very happy people. *Psychological Science, 13*, 81–84.

Digman, J. M. (1997). Higher order factors of the Big Five. *Journal of Personality and Social Psychology, 73*, 1246–1256.

Dignath, C., Buettner, G., & Langfeldt, H.-P. (2008). How can primary school students learn self-regulated learning strategies most effectively? A meta-analysis on self-regulation training programs. *Educational Research Review, 3*, 101–129.

Duckworth, A. L., Kirby, T., Tsukayama, E., Berstein, H., & Ericsson, K. (2010). Deliberate practice spells success: Why grittier competitors triumph at the National Spelling Bee. *Social Psychological and Personality Science, 2*, 174–181.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*, 1087–1101.

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (GRIT–S). *Journal of Personality Assessment, 91,* 166–174.

Dusek, J. B., Mergler, N. L., & Kermis, M. D. (1976). Attention, encoding, and information processing in low- and high-test-anxious children. *Child Development, 47*, 201–207.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41*, 1040–1048.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1*, 3–14.

Ehrler, D. J., Evans, J. G., & McGhee, R. L. (1999). Extending Big-Five theory into childhood: A preliminary investigation into the relationship between Big-Five personality traits and behavior problems in children. *Psychology in the Schools, 36*, 451–458.

Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review, 2*, 290–309.

Goff, M., & Ackerman, P. L. (1992). Personality-intelligence relations: Assessing typical intellectual engagement. *Journal of Educational Psychology, 84,* 537–552.

Gottfried, A. E. (1985). Academic intrinsic motivation in elementary and junior high school students. *Journal of Educational Psychology, 77*, 631–645.

Graham, S., & Golan, S. (1991). Motivational influences on cognition: Task involvement, ego involvement, and depth of information processing. *Journal of Educational Psychology, 83*, 187–194.

Haller, E. P., Child, D. A., & Walberg, H. J. (1988). Can comprehension be taught? A quantitative synthesis of "metacognitive" studies. *Educational Researcher, 17*, 5–8.

Hansford, B. C., & Hattie, J. A. (1982). The relationship between self and achievement/performance measures. *Review of Educational Research, 52*, 123–142.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research, 58*, 7–77.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology, 57*, 253–270.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26,* 107–129.

Jensen-Campbell, L. A., & Graziano, W. G. (2001). Agreeableness as a moderator of interpersonal conflict. *Journal of Personality, 69*, 323–362.

John, O. P., Caspi, A., Robins, R. W., Moffitt, T. E., & Stouthamer-Loeber, M. (1994). The "little five": Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Development, 65*, 160–178.

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology, 87*, 765–780.

Kashdan, T. B., & Silvia, P. J. (2009). Curiosity and interest: The benefits of thriving on novelty and challenge. *Oxford Handbook of Positive Psychology, 2*, 367–374.

Kim, K. H. (2005). Can only intelligent people be creative? *Journal of Secondary Gifted Education, 16*, 57–66.

Klafehn, J., Li, C., & Chiu, C. Y. (2013). To know or not to know, is that the question? Exploring the role and assessment of metacognition in cross-cultural contexts. *Journal of Cross-Cultural Psychology, 44*, 963–991.

Klein, H. J., & Lee, S. (2006). The effects of personality on learning: The mediating role of goal setting. *Human Performance, 19*, 43–66.

Koenig, J. A. (Ed.). (2011). *Assessing 21st century skills: Summary of a workshop*. Washington, DC: National Academies Press.

Komarraju, M., Karau, S. J., & Schmeck, R. R. (2009). Role of the Big Five personality traits in predicting college students' academic motivation and achievement. *Learning and Individual Differences, 19*, 47–52.

Kugle, C. L., Clements, R. O., & Powell, P. M. (1983). Level and stability of self-esteem in relation to academic behavior of second graders. *Journal of Personality and Social Psychology, 44*, 201–207.

Kuntsche, E., Knibbe, R., Gmel, G., & Engels, R. (2006). Who drinks and why? A review of socio-demographic, personality, and contextual issues behind the drinking motives in young people. *Addictive Behaviors, 31*, 1844–1857.

Kyllonen, P. C., Lipnevich, A. A., Burrus, J., & Roberts, R. D. (2014). *Personality, motivation, and college readiness: A prospectus for assessment and development* (ETS RR-14-06). Princeton, NJ: Educational Testing Service.

Laursen, B., Pulkkinen, L., & Adams, R. (2002). The antecedents and correlates of agreeableness in adulthood. *Developmental Psychology, 38*, 591–603.

Leary, M. R. (2005). The scientific study of personality. In V. Derlega, B. Winstead, & W. Jones (Eds.), *Personality: Contemporary theory and research* (3rd ed.). Chicago: Nelson-Hall.

Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *Journal of Educational Psychology, 97*, 184–196.

Liu, O. L., Rijmen, F., MacCann, C., & Roberts, R. (2009). The assessment of time management in middle-school students. *Personality and Individual Differences, 47*, 174–179.

Lloyd, J., & Barenblatt, L. (1984). Intrinsic intellectuality: Its relations to social class, intelligence, and achievement. *Journal of Personality and Social Psychology, 46,* 646–654.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice-Hall.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist, 57*, 705–717.

Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science, 15*, 265–268.

MacCann, C., Duckworth, A. L., & Roberts, R. D. (2009). Empirical identification of the major facets of conscientiousness. *Learning and Individual Differences, 19*, 451–458.
MacCann, C., Fogarty, G. J., Zeidner, M., & Roberts, R. D. (2011). Coping mediates the relationship between emotional intelligence (EI) and academic achievement. *Contemporary Educational Psychology, 36*, 60–70.
MacCann, C., Lipnevich, A., Burrus, J., & Roberts, R. (2012). The best years of our lives? Coping with stress predicts school grades, life satisfaction, and feelings about high school. *Learning and Individual Differences, 22*, 235–241.
Journal of Research in Personality. Retrieved: www.elsevier.com/locate/jrp
Marsh, H. W., Hau, K. T., Artelt, C., Baumert, J., & Peschar, J. L. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing, 6,* 311–360.
Maruyama, G., Rubin, R. A., & Kingsbury, G. G. (1981). Self-esteem and educational achievement: Independent constructs with a common cause? *Journal of Personality and Social Psychology, 40*, 962–975.
McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of Personality and Social Psychology, 52*, 1258–1265.
McCrae, R. R. (1996). Social consequences of experiential openness. *Psychological Bulletin, 120*, 323–337.
Mol, S. T., Born, M. P., Willemsen, M. E., & van der Molen, H. T. (2005). Predicting expatriate job performance for selection purposes: A quantitative review. *Journal of Cross-Cultural Psychology, 36*, 590–620.
Moutafi, J., Furnham, A., & Tsaousis, I. (2006). Is the relationship between intelligence and trait Neuroticism mediated by test anxiety? *Personality and Individual Differences, 40*, 587–597.
Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*, 30–38.
Nicholson, N., Soane, E., Fenton-O'Creevy, M., & Willman, P. (2005). Personality and domain-specific risk taking. *Journal of Risk Research, 8*, 157–176.
Nottelman, E. D., & Hill, K. T. (1977). Test anxiety and off-task behavior in evaluative situations. *Child Development, 48*, 225–231.
Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*, 543–578.
Partnership for 21st Century Skills. (2012). *Our history.* Retrieved from http://www.p21.org/about-us/our-history
Payton, J., Weissberg, R. P., Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Schellinger, K. B., & Pachan, M. (2008). *The Positive Impact of Social and Emotional Learning for Kindergarten to Eighth-Grade Students: Findings from Three Scientific Reviews.* Technical Report. Collaborative for Academic, Social, and Emotional Learning (NJ1).
Peabody, D., & de Raad, B. (2002). The substantive nature of psycholexical personality factors: A comparison across languages. *Journal of Personality and Social Psychology, 83*, 983–997.
Perugini, M., & Gallucci, M. (1997). A hierarchical faceted model of the Big Five. *European Journal of Personality, 11*, 279–301.
Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33–40.
Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*, 322–338.
Pottebaum, S. M., Keith, T. Z., & Ehly, S. W. (1986). Is there a causal relation between self-concept and academic achievement? *Journal of Educational Research, 79*, 140–144.
Preckel, F., Holling, H., & Vock, M. (2006). Academic underachievement: Relationship with cognitive motivation, achievement motivation, and conscientiousness. *Psychology in the Schools, 43*, 401–411.
Prins, P. J. M., Groot, M. J. M., & Hanewald, G. J. (1994). Cognition in test-anxious children: The role of on-task and copying cognition reconsidered. *Journal of Consulting and Clinical Psychology, 62*, 404–409.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*, 353–387.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychological Bulletin, 132,* 1–25.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*, 261–288.

Robert, R. D., Martin, J., & Olaru, G. (2015). *The Big Five personality factors as the Rosetta Stone for understanding, assessing, and enhancing noncognitive skills in K-12 education worldwide*. New York: Asia Society.

Roberts, B. W., Bogg, T., Walton, K. E., Chernyshenko, O. S., & Stark, S. E. (2004). A lexical investigation of the lower-order structure of conscientiousness. *Journal of Research in Personality, 38*, 164–178.

Roberts, B. W., Chernyshenko, O. S., Stark, S. E., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology, 58*, 103–139.

Roberts, R. D., Krause, H., & Suk-Lee, L. (2001). *Australian time organization and management scales*. Unpublished Inventory: University of Sydney.

Roberts, R. D., & Kyllonen, P. C. (1999). Morningness-eveningness and intelligence: Early to bed, early to rise will likely make you anything but wise. *Personality and Individual Differences, 27*, 1123–1133.

Robins, R. W., Tracy, J. L., Trzesniewski, K., Potter, J., & Gosling, S. (2001). Personality correlates of self-esteem. *Journal of Research in Personality, 35*, 463–482.

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal, 24*, 92–96.

Sadowski, C. J., & Gulgoz, S. (1992). Association of need for cognition and course performance. *Perceptual and Motor Skills, 74*, 498.

Saucier, G., & Ostendorf, F. (1999). Hierarchical subcomponents of the Big Five personality factors: A cross-language replication. *Journal of Personality and Social Psychology, 76*, 613–627.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology, 86*, 162–173.

Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist, 25*, 71–86.

Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research, 4,* 27–41.

Selfhout, M., Burk, W., Branje, S., Denissen, J., Van Aken, M., & Meeus, W. (2010). Emerging late adolescent friendship networks and Big Five personality traits: A social network approach. *Journal of Personality, 78*, 509–538.

Shute, V. J., Ventura, M., Kim, Y. J., & Wang, L. (2014). Video games and learning. In W. G. Tierney, Z. Corwin, T. Fullerton, & G. Ragusa (Eds.), *Postsecondary play: The role of games and social media in higher education* (pp. 217–235). Baltimore, MD: John Hopkins University Press.

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology, 100*, 330–348.

Stark, S., Chernyshenko, O. S., Drasgow, F., & White, L. A. (2012). Adaptive testing with multidimensional pairwise preference items: Improving the efficiency of personality and other noncognitive assessments. *Organizational Research Methods, 15*, 463–487.

Steel, P., Schmidt, J., & Shultz, J. (2008). Refining the relationship between personality and subjective well-being. *Psychological Bulletin, 134*, 138–161.

Tough, P. (2011). What if the secret to success is failure? *The New York Times Magazine.* Retrieved from http://www.nytimes.com/2011/09/18/magazine/what-if-the-secret-to-success-is-failure.html?_r=1&pagewanted=all

Tough, P. (2012). *How children succeed: Grit, curiosity, and the hidden power of character*. New York: Houghton Mifflin Harcourt Publishing.

Tupes, E. C, & Christal, R. E. (1961). *Recurrent personality factors based on trait ratings* (Tech. Rep. No. ASD-TR-61-97). Lackland Air Force Base, TX: U.S. Air Force.

Vollrath, M., & Torgersen, S. (2002). Who takes health risks? A probe into eight personality types. *Personality and Individual Differences, 32*, 1185–1197.

von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science, 6*, 574–588.

Wang, L., MacCann, C., Zhuang, X., Liu, O. L., & Roberts, R. D. (2009). Assessing teamwork and collaboration in high school students: A multimethod approach. *Canadian Journal of School Psychology, 24*, 108–124.

Woo, E. S., Harms, P. D., & Kuncel, N. R. (2007). Integrating personality and intelligence: Typical intellectual engagement and need for cognition. *Personality and Individual Differences, 43*, 1635–1639.

Wylie, R. C. (1979). *The self-concept: Vol. 2. Theory and research on selected topics*. Lincoln: University of Nebraska Press.

Zacharatos, A., Barling, J., & Kelloway, E. K. (2000). Development and effects of transformational leadership in adolescents. *The Leadership Quarterly, 11*, 211–226.

Ziegler, M., MacCann, C., & Roberts, R. (Eds.). (2011). *New perspectives on faking in personality assessment*. New York: Oxford University Press.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*, 3–17.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*, 166–183.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29*, 663–676.

Zimmerman, M. A., Copeland, L. A., Shope, J. T., & Dielman, T. E. (1997). A longitudinal study of self-esteem: Implications for adolescent development. *Journal of Youth and Adolescence, 26*, 117–141.

Chapter 2
Test Standards and Psychometric Modeling

Matthias Ziegler and Martin Brunner

2.1 Introduction

Psychosocial skills, abilities, and knowledge drive students' behavior and development in and outside of school. In recent years the importance of questionnaires assessing psychosocial skills has increased tremendously as researchers and practitioners realize the added value when combined with information derived from cognitive tests. Given the broad impact of students' cognitive and psychosocial skills, psychological assessment of students' individual characteristics is relevant for many purposes. Decisions pertaining to the cognitive ability of a student and their fitness to stand trial, the need for special schooling, or therapy are just a few examples. Obviously such decisions bare great relevance not only for the student assessed but also for family and friends and society in general. Decisions with such a potentially widespread impact need to be based on solid data gathered with sound measures.

Psychological tests or questionnaires are often considered to be sound measures. However, caution is warranted. Not every test or questionnaire can be considered sound, nor do all measures represent a solid basis for informed decisions that may affect people's lives. Especially the assessment of psychosocial skills faces problems that are of less importance for the assessment of cognitive abilities. Socially desirable responding, faking or low convergence between different measures are just a few examples for such issues which we will outline in this chapter. The reason

M. Ziegler (✉)
Psychological Institute, Humboldt-Universität zu Berlin,
Unter den Linden 6, Berlin, Germany
e-mail: zieglema@hu-berlin.de

M. Brunner
Department of Education and Psychology, Free University of Berlin,
Otto-von-Simson-Straße 15, Berlin, Germany
e-mail: martin.brunner@isq-bb.de

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_2

for the greater susceptibility of psychosocial skill measures to these issues can be found in their operationalization. Whereas answers to items in cognitive tests can clearly be classified as right or wrong, such a classification is problematic for self- and other ratings, which are most often used to assess psychosocial skills. Cattell (1958) used the term Q-data (questionnaire data) and T-data (test data) to differentiate between subjective questionnaires and objective tests. Objective here refers to the possibility to clearly define an answer as right or wrong. Not being able to do this with questionnaires brings along many problems. Despite this, it has to be noted that such questionnaires have been shown to be useful in predicting a wide range of real-life criteria (Barrick & Mount, 1991; Lindqvist & Vestman, 2011; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007).

The overarching goal of the present chapter is therefore to provide guidance to users but also developers of psychological measures to help selecting, administering, and constructing psychological assessment instruments. To this end, we draw on the test standards as theoretical framework to address general principles (e.g., reliability and validity) that are important in the context of psychological assessment in Sect. 2.1. In Sect. 2.2, we elaborate on more complex psychometric models and show how these models are linked to test standards, specifically for computing and understanding score reliability. In Sect. 2.3, we conclude by discussing some of the main challenges facing psychosocial assessment with a focus on response sets and response styles as well as general problems with test-criterion correlations.

2.2 Psychological Assessment: Process and Standards

2.2.1 Psychological Assessment as Process

Psychological assessment usually starts with a more or less straightforward question such as “Does Mary require special schooling?” The question obviously entails a number of related questions that need to be answered. Moreover, the person working to answer the question needs to be an expert in the field to understand what is required to come up with a trustworthy answer. Thus, one of the first steps in any psychological assessment process is the systematic delineation of sub-questions, otherwise known as specific hypotheses. Specific hypotheses often cover areas such as cognitive ability, knowledge, personality, social network, or skills. Once specific hypotheses have been formulated, it is necessary to devise means to gather data that help to answer them. This is called operationalization. In many cases psychological tests such as tests of general mental ability or questionnaires assessing an individual’s personality can be helpful tools to operationalize a hypothesis and gather data. Methods like behavior observation and interviewing should not be forgotten, though. While this all sounds rather straightforward, practitioners know that the market for assessment instruments is large and at times confusing. To assess a certain construct, practitioners can choose between a plethora of instruments, whose

exact number is hard to grasp. Ziegler, Booth, and Bensch (2013) showed that scores from inventories supposedly capturing the same construct do not necessarily do so. In other words, two questionnaires that purport to assess Extraversion do not necessarily come to the same result simply because the underlying construct actually measured is slightly different across the two measures. This may seem like a rare exception. However, there is a wide range of literature highlighting the problem (e.g., Miller, Gaughan, Maples, & Price, 2011; Mussel, 2010; Pace & Brannick, 2010). Thus, caution is warranted when selecting tests or questionnaires[1] mentioned above; this problem seems even more pronounced for measures assessing psychosocial skills. Here construct definitions are often more fuzzy, and the delineation of items is therefore more arbitrary than for measures assessing cognitive abilities. Such abilities are often clearly defined leaving little leeway when constructing items. As a consequence, the measures developed to capture the same psychosocial skill might contain items with seemingly diverging content. As a result, a person's score in such tests might differ. Thus, before selecting a test or before formulating items for a test, it is essential to clearly define the construct(s) intended to be measured. Such a definition should not only contain examples for how the construct manifests itself in behavior, it should also specify which constructs are closely related and which constructs are not intended to be measured. That way, the nomological network of the construct to be measured is specified.

Within the domains of psychology and education, there are widely accepted standards for testing that have been developed in a joint partnership by the American Psychological Association (APA), the American Educational Research Association (AERA), and the National Council on Measurement in Education (NCME). These Standards for Educational and Psychological Testing (or simply *the Standards*) are seen by many as the authoritative source for the development and evaluation of educational and psychological tests and assessments. We will use the Standards as our guiding framework in the following sections. Practitioners looking for operationalizations to answer their specific hypotheses will find the Standards to be a helpful companion when selecting tests, evaluating existing test results, or constructing tests.

2.2.2 The Standards

The latest version of the Standards was published in 1999. This was the fourth version of the Standards jointly published by APA, AERA, and NCME since 1966. A joint committee (see http://teststandards.org/Revision.htm) has been working on another revision since 2009. This new version was announced to be published in spring 2014 but has not been released until now (i.e., August 2014). Thus, we will

[1] The remainder of this chapter mainly revolves around tests and questionnaires. It is important to note, though, that the following quality issues also apply to interviews and behavior observations!

adhere to definitions and guidelines from the 1999 version but try to outline expected changes as well.

The Standards (1999 version) consist of three parts: (I) Test Construction, Evaluation, and Documentation; (II) Fairness in Testing; and (III) Testing Applications. For the new 2014 version, the APA website gives the following information: "The overall organization of the revised Standards is also different from that of the 1999 edition. The new version is separated into "Foundations," "Operations" and "Testing Applications" sections. The "Foundations" section focuses on fundamental testing issues such as validity, reliability, and fairness. The "Operations" section deals with operational testing issues such as test design and development, test administration, scoring and reporting, and supporting documentation for tests. The "Testing Applications" section details specific applications in testing such as workplace testing and credentialing, educational testing and assessment, and the use of tests for program evaluation, policy studies, and accountability." Within the psychoeducational area, guidelines regarding test applications are especially important. However, the best test application is pointless, if the test chosen is problematic to begin with. Consequently, great care is warranted when selecting or constructing tests. Moreover, a test result is only as good as the report it is used in. We will therefore focus on reliability, validity, norms, reporting, and fairness. These standards are especially important when operationalizing specific hypotheses within a diagnostic process as outlined above.

Before delving into the details of psychometric standards, it is important to note that reliability and validity always refer to the interpretation of a test score in a certain setting. This means that a test is neither reliable nor valid per se. Rather than that, the interpretation of a test score derived from administering the test under certain circumstances can be reliable and valid.

Reliability and Measurement Error To explain the concept of reliability, we draw on classical test theory, CTT (Lord & Novick, 1968). One of the main ideas underlying CTT is that each person has a so-called true score on the latent variable (or psychological construct) to be measured. However, since the variable to be measured is latent, one needs to observe manifestations of the construct in question. In this sense, one cannot just go ahead and assess, for example, the latent construct Life Satisfaction. Instead, measures of Life Satisfaction ask respondents to report how satisfied they are with various aspects of their lives. These responses can then be observed and scored. Moreover, using just one measurement point (i.e., one item) usually does not suffice. Due to measurement errors, the observed score does not match the true score exactly. However, using aggregation across many measurement occasions (items), the measurement error hypothetically becomes smaller and smaller (reasons for this are explained below).

Thus, CTT assumes that an observed score (X) equals the sum of a person's true score (T) and measurement error (E):

$$X = T + E \tag{2.1}$$

This is one of the fundamental axioms of CTT (also see Sect. 2.1). If this axiom is generalized across a population, one can say that the variance of the observed scores $\left(\hat{\sigma}_x^2\right)$ equals the sum of the true score variance $\left(\hat{\sigma}_T^2\right)$ and the error variance $\left(\hat{\sigma}_E^2\right)$:

$$\hat{\sigma}_x^2 = \hat{\sigma}_T^2 + \hat{\sigma}_E^2 \tag{2.2}$$

This general equation is only true if a few other assumptions of CTT hold. For example, the error must be random which means it does not correlate with errors in other tests, the true score of the test, or true scores in other tests. Under these, admittedly very optimistic, boundaries, reliability (Rel) refers to the amount of true score variance (or systematic variance) within the observed variance:

$$\text{Rel} = \frac{\hat{\sigma}_T^2}{\hat{\sigma}_x^2} \tag{2.3}$$

Within the framework of CTT, score reliability is thus mathematically defined in terms of the proportion of true score variance to observed score variance (Lord & Novick, 1968, p. 61). When aggregating measures that fulfill these assumptions, the relative amount of true score variance increases in every aggregation step because the random error cannot add up across the different measures whereas the systematic true score variance can. Reliability may range between 0 (no reliability) and 1 (perfect reliability). Conceptually, reliability can thus be considered as an index of measurement precision (Lord & Novick, 1968; McDonald, 1999; Mellenbergh, 1996). A less mathematical definition of reliability is: reliability is the degree of accuracy a test score has. The 1999 Standards state:

> A test, broadly defined, is a set of tasks designed to elicit or a scale to describe examinee behavior in a specified domain, or a system for collecting samples of an individual's work in a particular area. Coupled with the device is a scoring procedure that enables the examiner to quantify, evaluate, and interpret the behavior or work samples. Reliability refers to the consistency of such measurements when the testing procedure is repeated on a population of individuals or groups. (p. 25)

The key problem to compute reliability of test scores is that we just do not know the true scores and the measurement errors, respectively. Many psychometricians have worked on quantifying the extent of the measurement error. As a result different ways of estimating reliability have been suggested. Most of these ways have in common that they use the aggregation principle (i.e., a score becomes more reliable if it is based on more than one measurement occasion). The two most prominent conceptions of reliability are internal consistency and test-retest reliability.

Cronbach's alpha (Cronbach, 1951), a measure of internal consistency, is currently probably the most commonly used reliability estimate. Below we will outline McDonald's (1999) approach to reliability, which is generally a much less error-prone estimate (see Sects. 2.2 and 2.3). There are many other ways to estimate a test

score's reliability, for example, parallel test reliability (Cronbach, 1947), split-half reliability (Kuder & Richardson, 1937), and construct reliability (see Sect. 2.2). There are also other theoretical frameworks like generalizability theory (Rajaratnam, Cronbach, & Gleser, 1965; Shavelson & Webb, 2006) or item response theory (Embretson & Reise, 2000). The reader is encouraged to familiarize with all of these to round up the rough sketch provided here.

The idea underlying Cronbach's alpha is that each test item should order the test takers according to their true score. As a result of this, the items will correlate. The higher the correlation, the larger the agreement in ordering will be. In fact, it is assumed that all deviations from a perfect ordering across items are due to random measurement error. Thus, reliability in terms of Cronbach's alpha is a function of the number of homogeneous items (i.e., items measuring the same latent construct) and the extent of item overlap (i.e., inter-item correlations). This makes the interpretation of Cronbach's alpha a bit difficult. Imagine a rather broad construct such as Extraversion. There are many different behaviors that can be regarded as manifestations of Extraversion, for example, being talkative, being sociable, being assertive, being adventure seeking, and so on. Thus, the item overlap will be moderate at best, and in order to cover the whole construct of Extraversion—that is, to have a valid measure of Extraversion—many different items are needed. Extraversion measures using only a few items but yielding scores with large Cronbach's alpha estimates should therefore be considered with caution. The question is whether such measures still cover the whole nomological net of Extraversion. In other words, before judging a Cronbach's alpha, it is necessary to consider the breadth of the construct to be assessed in relation to the number of items. Without going into detail an example should make this clearer. A questionnaire assessing Extraversion and consisting of 48 items could have a Cronbach's alpha that is comparable in size to another Extraversion measure with only four items. Now, both measures might have the same internal consistency; however, their content and thus the breadth of construct coverage most likely are very different.

Test-retest reliability also uses the aggregation principle. Using this approach, aggregation is achieved by testing twice with the same instrument. There should be an adequate (in terms of construct measured and measurement goal) time gap between both measurement points. Again, if each measurement occasion is used to order test takers according to their true scores, there should be a perfect rank order agreement (i.e., a perfect correlation). Deviations from this perfect ordering are regarded as the influence of a random measurement error. In other words, the extent of agreement is a direct estimate of the shared systematic and therefore true score variance. Of course, differential learning or repetition effects need to be considered because they will cause changes in rank ordering as well.

Depending on the specific hypotheses that need to be answered in a psychological assessment process a specific reliability estimate is required. Status assessments, i.e., assessments regarding the status quo of a person, usually require estimates like Cronbach's alpha or construct reliability (see Sect. 2.2). Prognoses, on the other hand, where the level on a target construct is to be predicted across time require test-retest reliabilities. Thus, a test manual containing important test information

should include detailed descriptions regarding the aimed use of the instrument, the construct meant to be measured, the type of reliability estimated, the circumstances of data collection, and so forth. The 1999 Standards provide important guidelines: "For each total score, subscore, or combination of scores that is to be interpreted, estimates of relevant reliabilities and standard errors of measurement or test information functions should be reported" (p. 31).

It is important to note that a test can be reliable but does not really measure the target construct. For example, if we used the length of a student's forehead multiplied by his or her shoe size, we would have a very reliable score as these quantities change only marginally in a certain period of time. If we said that this score assesses students' Life Satisfaction, we would clearly be off target. Thus, the score would be reliable but not valid.

Validity The 1999 Standards define validity as:

> [T]he degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests. Validity is, therefore, the most fundamental consideration in developing and evaluating tests. The process of validation involves accumulating evidence to provide a sound scientific basis for the proposed score interpretations. It is the interpretations of test scores required by proposed uses that are evaluated, not the test itself. When test scores are used or interpreted in more than one way, each intended interpretation must be validated. (p. 9)

As was the case with reliability, there is not a single, right way to demonstrate the validity of test score interpretations. Accordingly, there are different types of validity evidence defined in the Standards: evidence based on (I) content, (II) response process, (III) internal structure, (IV) relations with other variables, and (V) consequences of testing.

Content Validity Evidence based on content is also called *content validity* and refers to the degree with which the test content reflects the construct intended to measure. To judge this evidence one needs to examine the "themes, wording, and format of the items, tasks, or questions" according to the Standards (p. 11). In most cases this type of validity evidence is derived from expert ratings.

Evidence from Response Processes The idea underlying *evidence based on response processes* is rather straightforward. If a measure is supposed to capture the construct Life Satisfaction, individuals should be asked to judge their satisfaction with various aspects of their lives. Analyses of the response process can be conducted with think-aloud techniques or retrospective interviews (Krosnick, 1999). Especially during item evaluation such techniques can provide useful information (e.g., whether persons understand the contents of an item in the way it is supposed to be understood).

Internal Structure Evidence based on internal structure is also called *factorial validity*. The idea is that items assessing a certain construct demonstrate a hypothesized structure. This means that ideas exist regarding the number of dimensions a measure has (e.g., for measures of subjective well-being, one should be able to

distinguish the dimension Life Satisfaction from dimensions reflecting positive and negative affect), the allocation of items to these dimensions, and possibly a hierarchical structure describing the relations between the dimensions. Evidence for the internal structure can be provided using techniques such as confirmatory factor analysis. It is important though that the structure tested is congruent with the theoretical model described by the test authors. Another important issue directly related to this is the use of correlated residuals within confirmatory factor analyses. In such analyses, item variance is decomposed into a systematic trait component and a residual, i.e., variance not explained by the trait intended to be measured. In order to evaluate such models, different criteria are used (e.g., Heene, Hilbert, Draxler, Ziegler, & Bühner, 2011; Heene, Hilbert, Freudenthaler, & Bühner, 2012). Oftentimes, models fail to satisfy these criteria and are therefore changed. This in itself is problematic for factorial validity as an a priori model should be tested and not changed. Thus, any change requires a new testing using new data. Despite this, a common way to improve model fit is to allow correlations between residuals. This basically means that items still share systematic variance once the construct intended to be measured is controlled for. While this might simply be variance due to method effects (e.g., both items start with the phrase "I like to…"), it might also imply that the items in question are not unidimensional, i.e., they assess an additional trait. This would be a serious problem for assessment. For example, if a test was supposed to measure Life Satisfaction but some of the items have correlated errors because they also assess Extraversion, it would be unclear if an especially high or low score in the test was due to Life Satisfaction or Extraversion. Consequently, when judging factorial validity evidence, it is important to look carefully at the model tested.

Relations with Other Variables Evidence based on relations with other variables can also be considered as *construct validity*. In that sense, evidence can be provided if a test score correlates with a test score from another test capturing a similar construct (*convergent validity*) but does not substantially correlate (or correlates considerable less) with a test score from a test capturing a different construct (*discriminant validity*). These ideas go back to the seminal paper by Campbell and Fiske (1959) outlining a method used to display convergent and discriminant validity (i.e., the multitrait-multimethod matrix). Modern approaches still use the same general ideas but apply more stringent ways of testing, for example, using structural equation modeling (Eid et al., 2008). In that sense, the variance of test scores is separated into trait variances and method variance. For example, if Life Satisfaction and Extraversion are measured by self- and by other reports, multiple traits (two) have been assessed with multiple methods (two). Using structural equation modeling, it is then possible to estimate the variance within a test score due to the respective trait (i.e., Life Satisfaction or Extraversion) and the method (i.e., self- or other reports). At the same time, different components of the model can be interpreted as evidence for convergent validity (e.g., trait loadings of measures capturing the same trait across methods) and discriminant validity (e.g., the correlation between traits).

Oftentimes test scores are supposed to predict a specific behavior or feeling. Therefore, it is necessary that the interpretation of a test score can be generalized to everyday situations, outside of the test situation. Validity evidence for this requirement is called *test criterion validity* and also falls within the category of evidence based on relation with other variables. This method usually calls for a correlation between a test score and a relevant criterion. Depending on several other factors (e.g., range restriction, reliability of test score and criterion, level of symmetry between test score and criterion, criterion contamination, and criterion deficiency), the size of the actual test-criterion correlation needs to be considered carefully.

Factors Influencing Relations with Other Variables To exemplify these concepts, let us assume that Life Satisfaction was assessed in a group of students attending a summer sports camp for successful high school athletes. To this end, a questionnaire was administered which captures general Life Satisfaction using ten items. In order to determine a test-criterion correlation, students' scholastic performance was also measured. This was operationalized with a math test all students took at the end of the camp. This group of students certainly is highly selective. Not everyone is a sports star and can attend such a camp. Thus, it could be assumed that the differences in Life Satisfaction between such athletes are smaller than in a representative sample of high school students. If a test-criterion correlation is now estimated, the result might be lower than in a representative sample. The reason for this lies within the equation needed to estimate a correlation. Within the nominator, the covariance between test and criterion is calculated as the product sum between students' deviation from the Life Satisfaction mean of the sample times the students' deviation from the math test mean of the sample. If the variance of Life Satisfaction is restricted in range, students will have smaller deviations from the mean than students from a representative sample. As a consequence, the whole correlation could be smaller.

Likewise the reliability of the Life Satisfaction measure or the math test might lower the correlation. Above we have stated that reliability represents the amount of true score variance within a test. A correlation can only be derived from such systematic true score variance. Thus, tests with low reliability can only achieve comparably low correlations with other scores. This is known as the attenuation paradox (Loevinger, 1954) and can be solved using correction formulas (Fan, 2003). Thus, if either the Life Satisfaction measure or the math test or both have low reliabilities, their correlation will be lowered by this.

Another interesting phenomenon is the level of symmetry (Brunswik, 1955). The Life Satisfaction measure contains ten items and assesses a very general and abstract construct. The math test, however, is just one measure, assessing performance at a specific point in time and most likely in only a few mathematical areas. Thus, this performance is rather specific. Obviously, the levels on which test and criterion are measured differed. It would have been more appropriate to either use a more specific Life Satisfaction measure (e.g., Life Satisfaction experienced at school or school satisfaction) or a more general performance test. In general, the levels of measurement should be comparable for test and criterion.

Finally, the operationalization of the criterion can also influence the correlation. It was intended to measure scholastic performance. Math performance certainly is part of scholastic performance. However, math performance might also be influenced by factors that do not reflect the students' aptitude. For example, the quality of the math teacher certainly influences the students' performance. However, the questionnaire was not meant to predict this aspect of students' performance. On the other hand, scholastic performance is much more than just math. In that way, the operationalization of the criterion is both contaminated and deficient (Brogden & Taylor, 1950).

Consequences of Testing The idea behind evidence based on consequences of testing is that the concept of validity needs to include the intended and unintended consequences of testing. The 1999 Standards provide a very good example for this idea: "In the case of employment testing, if a test publisher claims that use of the test will result in reduced employee training costs, improved workforce efficiency, or some other benefit, then the validation would be informed by evidence in support of that claim" (pp. 16–17).

Final Thoughts As can be seen from the many different approaches to validity evidence, judging the validity of a test score interpretation is not an easy task. It gets easier with every piece of information provided by the test developer (or publisher). Accordingly, standard 1.1 in the 1999 Standards say: "A rationale should be presented for each recommended interpretation and use of test scores, together with a comprehensive summary of the evidence and theory bearing on the intended use or interpretation" (p. 17).

Before coming to an end with regard to validity, it should be mentioned that there are psychometricians who propose a different concept of validity. Borsboom, Mellenbergh, and Van Heerden (2004) argue: "A test is valid for measuring an attribute if and only if (a) the attribute exists and (b) variations in the attribute causally produce variations in the outcomes of the measurement procedure" (p. 1061). These authors do not refer to the validity of a test score interpretation. Moreover, the methods used to test for validity rely much stronger on the knowledge of processes that lead to item responses that should be found in substantive psychological theory.

Norms When taking a driver's license test, everyone has to fill out a test, and only those with less than a certain amount of errors are granted a license. This is a criterion-oriented test score interpretation. Many psychological tests do not work according to such a principle but rather use a norm-referenced test score interpretation. Such an interpretation determines an individual's standing by comparing it to a relevant group. In that sense, it is not important how many errors one makes in total. It is more important how the number of errors relates to the number of errors made in the reference group. Norm values like the intelligence quotient (IQ) work according to this principle. For example, the mean score for an intelligence test determined in a reference group is set to 100 and the standard deviation to 15. It is further assumed that intelligence test scores in a representative sample follow a

normal distribution. Therefore, between plus and minus one standard deviation from the mean, approximately two thirds of all test scores will lie. This area is often referred to as average. The remaining one third of scores are symmetrically distributed right and left from the average scores. Thus, approximately 16 % of scores are above and below average, respectively. Thus, if an individual gets assigned an IQ of 125 within the test, this means that the person's score is more than one standard deviation above the mean and belongs to the 16 % of people who are above average. This example should show that a test score based on a reference norm is basically useless information unless the norm group is specified. Therefore, when looking for the right test to operationalize a specific hypothesis, it is important to look at the information the test publisher provides for the norm group (e.g., age, gender composition, educational level, circumstances of assessment, etc.). If the wrong norm group is selected, the validity of the test score interpretation can be seriously endangered. Just imagine, a person's Life Satisfaction is estimated in reference to a norm group of people who have just been informed to have won the lottery instead of a norm group resembling the average population. The 1999 Standards write in 4.5 (p. 55): "Norms, if used, should refer to clearly described populations. These populations should include individuals or groups to whom test users will ordinarily wish to compare their own examinees." And in 4.6: " Reports of norming studies should include precise specification of the population that was sampled, sampling procedures and participation rates, any weighting of the sample, the dates of testing, and descriptive statistics. The information provided should be sufficient to enable users to judge the appropriateness of the norms for interpreting the scores of local examinees. Technical documentation should indicate the precision of the norms themselves."

Reporting The results of psychological testing are often reported to a variety of people with differing expertise in such matters. Consequently, the information needs to be wrapped up in a way that is easy to understand and makes sure that misinterpretations do not occur. A general problem in this regard is that reporting specific values (e.g., the test taker has an IQ of 111) suggests that the test score interpretation is perfectly reliable. Even though reliability is often rather good, it is mostly far from being perfect. Questionnaires assessing psychosocial constructs sometimes have rather low reliabilities (mostly internal consistencies). This is especially often the case if such measures capture lower-order personality facets. Now, one could argue that such facets with low reliability estimates should just not be used. However, personality facets (not only the unreliable ones) have been shown to have good test-criterion correlations (e.g., Paunonen & Ashton, 2001). Thus, reports should account for the issue of unreliability by using methods that take measurement error into consideration. One such method is the confidence interval (CI). CIs use the standard error of measurement, an unreliability estimate, and an accepted margin of error to estimate a range that covers the true score of a person with a certain probability. As a result a report does not contain a specific number like an IQ of 111 or a Life Satisfaction of T = 34. Rather than that an interval is reported based on the CI, e.g., IQ average (95 % CI 108–114) or Life Satisfaction below

average (80 % CI 30–38). Thus, using CIs when reporting test results should be a matter of course. The Standards (1999) also say:

> When test score information is released to students, parents, legal representatives, teachers, clients, or the media, those responsible for testing programs should provide appropriate interpretations. The interpretations should describe in simple language what the test covers, what scores mean, the precision of the scores, common misinterpretations of test scores, and how scores will be used. (p. 65)

Fairness Fairness in testing is an important issue, and test publishers need to ensure that appropriate studies are reported outlining this issue for the test presented. Generally, a test is fair if it does not systematically disadvantage test takers from specific subpopulations. In a more specific sense, four conceptions of fairness can be differentiated. *Fairness as lack of bias* means that test scores have the same meaning for members from different subgroups. Methods such as differential item functioning are used to test this (Stark, Chernyshenko, & Drasgow, 2006; Walker, 2011). *Fairness as equitable treatment in the testing process* means that every test taker should have comparable if not the same opportunity to demonstrate his or her standing on the construct measured. *Fairness as equality in outcomes of testing* requires that passing rates in a test should not be different among subgroups. *Fairness as opportunity to learn* becomes important when success in a test partly depends on the learning opportunities a test taker has had. If for example an achievement test score is low because the test taker did not receive adequate schooling, decisions based on the score would be unfair if the test taker had been better with more schooling. An example for a psychosocial test is the big-fish-little-pond effect on academic self-concept (Marsh, 1987). The idea here is that students use their class or school as a frame of reference when forming their academic self-concepts. In doing so, a student's level of academic self-concept depends on the average achievement level of his or her class or school. Thus, when two students have the same achievement, the student attending the class with a higher average level of achievement has a lower academic self-concept.

Summing up, the psychological assessment process requires the selection or construction of psychological tests. The Standards provide a framework for the quality of such tools with an emphasis on validity and reliability. The short descriptions above are supposed to give a first insight and make clear that no easy cookbook-like rules should be applied when judging reliability and validity. Particularly, this judgment should always take the construct itself, the method applied, the (number of) items, the norm group, and its specifics into account.

2.3 Test Standards, Score Reliability, and Psychometric Models

In general, psychosocial constructs are not directly observable but need to be inferred from observable behaviors (i.e., item answers). Psychometric models link latent constructs with observable measures and are therefore essential to make

informed judgments when selecting or constructing psychological tests. Specifically, many test standards are closely tied to the application of psychometric models because psychological constructs, in general, and psychosocial skills in particular are (unobserved) latent variables that are reflected in observable scores on corresponding measures (i.e., operationalization; see above).[2] One key task for educational and psychological researchers is therefore to choose from a variety of psychometric models that link latent constructs with observable measures. Importantly, it is this decision that enables researchers to empirically tackle questions whether the scores of their measures fulfill test standards. For example, when the psychometric model approximates the empirical data well, this finding provides a statistical rationale for the computation of scale scores that estimate respondents' levels on (latent) psychosocial constructs. Such a rationale is a prerequisite of the 1999 Standards, which state that "where composite scores are developed, the basis and rationale for arriving at the composites should be given" and that "the rationale and supporting evidence must pertain directly to the specific score […] to be interpreted or used" (p. 20). Further, when the psychometric model fits the data well, this can be considered as a crucial prerequisite for computing many statistical coefficients that assess score reliability (Cortina, 1993; Slaney & Maraun, 2008).

Taken together, psychometric models are of critical importance for both research and applied assessment. In the remainder of this section, we provide an in-depth discussion of the properties of confirmatory factor analysis (CFA) because CFA is a widely used and highly versatile psychometric model to tackle substantive questions on psychosocial skills. Further, we elaborate on the implications of CFA models for the computation and interpretation of model-based estimates of score reliability. Relative to widely used coefficients such as Cronbach's alpha (Cronbach, 1951), model-based estimates of reliability have the advantage that these reliability coefficients can be directly estimated from the CFA model applied. We believe that this discussion is informative to many researchers because in many cases the link between the assumed psychometric model and the reported reliability coefficient such as Cronbach's alpha is not made explicit.

2.3.1 Confirmatory Factor Analysis

To gather reliable and valid information about persons' psychosocial skills, individuals respond to measures, such as items on tests or questionnaires. When the data format of the measures is continuous or is at least approximately continuous, (confirmatory) factor analysis is a versatile psychometric model. In the present chapter we consider the factor model as it is defined for confirmatory factor analysis. The key idea of CFA is to relate individuals' levels on the latent construct(s) θ (which are also called latent variable or factor) to the observed measures X by means of a

[2] This section borrows from the article written by Brunner, Nagy, and Wilhelm (2012).

linear link function. The factor model for the score X_{ij} on a single measure j for person i can be written as:

$$X_{ji} = \lambda_j \cdot \theta_i + e_{ji} \tag{2.4}$$

Here λ_j is the standardized factor loading of manifest measure X_j on the latent target construct.

We illustrate the one-factor CFA model for the Satisfaction with Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985)—a widely used measure of Life Satisfaction. More specifically, the five-item scores of the SWLS are supposed to measure the latent construct *Life Satisfaction*. Data stem from the Luxembourgish extension of the year 2009 cycle of the PISA study (for study details, see Spengler, Lüdtke, Martin, & Brunner, 2013) where some 2500 students (attending grade 9 or 10) used a six-point rating scale (ranging from 1 = "strongly disagree" to 6 = "strongly agree") to answer the items of the German translation of the SWLS (Schumacher, Klaiberg, & Brähler, 2003). The CFA equations for the five-item scores (SWLS 1 to SWLS 5) of student i can be written as (Fig. 2.1a shows the graph for this functional relation for the first item "In most ways my life is close to my ideal"):

$$\begin{aligned}
\text{SWLS1}_i &= 0.81 \times \textit{Life Satisfaction}_i + e_{1i} \\
\text{SWLS2}_i &= 0.89 \times \text{Life Satisfaction}_i + e_{2i} \\
\text{SWLS3}_i &= 0.85 \times \text{Life Satisfaction}_i + e_{3i} \\
\text{SWLS4}_i &= 0.82 \times \textit{Life Satisfaction}_i + e_{4i} \\
\text{SWLS5}_i &= 0.82 \times \textit{Life Satisfaction}_i + e_{5i}
\end{aligned} \tag{2.5}$$

A few characteristics of the CFA model in Eq. 2.5 merit further consideration. First, the one-factor model implies that higher scores on *Life Satisfaction* are associ-

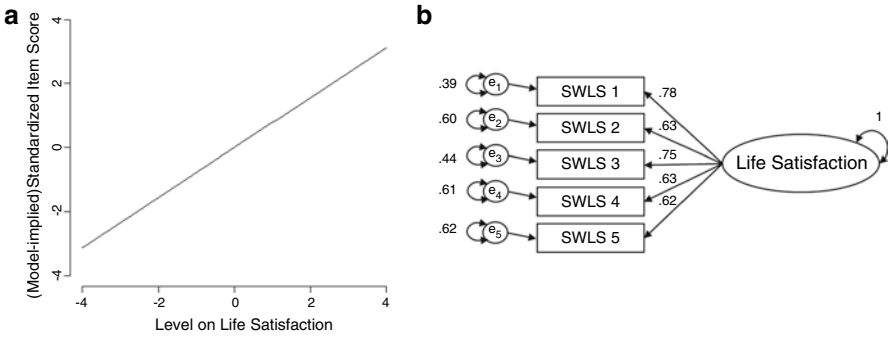

Fig. 2.1 Confirmatory factor analysis model: (**a**) Functional relation of the latent construct *Life Satisfaction* to the observed score on the item SWLS 1 "In most ways my life is close to my ideal" and (**b**) corresponding structural model for the *Life Satisfaction* construct as measured by the five items of the Satisfaction with Life Scale (SWLS). Standardized model parameters are shown. Model fit: $\chi^2(5, N = 2581) = 62.2$, $p < 0.001$; *CFI* = 0.98; *RMSEA* = 0.07; *SRMR* = 0.03

ated with higher scores on all five items. Thus, when the one-factor model fits well to the data, the latent construct *Life Satisfaction* accounts for the common variance across measures. In other words, when the model fits well, *Life Satisfaction* explains the intercorrelations among item scores.

Second, in the one-factor model, each item score is also affected by a second factor orthogonal to *Life Satisfaction*. This second factor may represent some item-specific aspect that affects individuals' responses to that item (e.g., individual differences in response sets; see Sect. 4.1). Further, each item score may also be affected to some degree by random measurement error. Both of these latter influences (i.e., reliable but item-specific variance and unreliable error variance) are represented by a single factor e_j for each item. Note that the factor e_j is specific to each item for two reasons: (a) Unless two measures share measure-specific variance (e.g., when the same item is applied at two successive points of measurement or when two self-report items have similar wordings), it is not possible to disentangle the variance of a particular subtest that is attributable to random measurement error from that attributable to specific variance. (b) Measurement error is uncorrelated across items because its influence on items is random (i.e., unpredictable according to the axioms of CTT; see above). Thus, the factors e_1 to e_5 as well as *Life Satisfaction* are supposed to operate mutually independently; these factors are therefore specified to be mutually uncorrelated.

Third, the CFA model is often depicted in terms of a structural model (see Fig.2.1b) where observed measures (depicted as rectangles) are linked to a latent construct (depicted as ellipses or circles) by a unidirectional arrow. Likewise, item-specific factors (i.e., e_1 to e_5) are depicted as circles pointing to the individual items in Fig.2.1b. In doing so, the structural diagram represents the idea that individual differences in the latent construct (where the double-headed arrow represents the variance of a latent construct) cause individual differences in the manifest measures.

Fourth, to simplify the CFA model, Eq. 2.5 does not include a parameter for the intercept of the measure. Moreover, to ease interpretation of model parameters (for the purposes of the present chapter), it is also assumed that the manifest item scores, and the latent factors *Life Satisfaction* and e_1 to e_5, are z-standardized (with $M = 0$ and $SD = 1$), respectively. Equation 2.5 implies, for example, that individuals with a level of *Life Satisfaction* of one standard deviation above the mean are expected to have a score that lies $0.78 * 1 = 0.78$ standard deviations above the mean on the item SWLS 1. Further, depending on the research goal, the CFA model can also easily be extended to include measure-specific intercepts, for example, when the CFA model is used to address questions on the fairness/invariance of measures across groups (e.g., Vandenberg & Lance, 2000) or time (e.g., McArdle, 2009).

Fifth, Eq. 2.4 shows that the score on a certain measure X_j is decomposed into one part that is linked to the latent target construct θ and another part that is linked to a (latent) residual term e_j. Thus, the CFA model can also be used to represent the axioms of CTT (see above and Bollen, 1989) in which a person's observed score X_i is considered to be composed of a person's true score T_i and a person-specific random error component E_i. However, the CFA model is more flexible than CTT as the

residual term e_{ji} in Eq. 2.1 may comprise both reliable variance that is specific to a certain measure and random error variance. This should be kept in mind, when CFA is used to model longitudinal data where the same measures are applied several times and where residual terms of the same measures are often found to be correlated across time (e.g., Hoyle & Smith, 1994). This model specification would not be possible within the framework of CTT because measurement error is assumed to be uncorrelated across measures with the influence of measurement error on subtests being truly random (i.e., unpredictable).

2.3.2 Psychometric Models and Reliability

Psychosocial skills are not directly observable entities, but latent variables. To assess an individual's psychosocial skills, we have to estimate his or her level on the respective latent variable. In most applied psychological research, several observed scale indicators are summed using unit weights (i.e., each scale indicator has the same weight in the computation of the sum score) to form a manifest scale score. This scale score gives an estimate of the person's level on the latent construct (see Grice, 2001). For example, when a person has completed the SWLS, a scale score reflecting his or her level of *Life Satisfaction* can be computed by using unit weights to sum up his or her scores on the five items of the SWLS to obtain the scale score "Life Satisfaction." But how reliable is this scale score?

To answer this question, we have to remember how reliability can be mathematically defined. Within the framework of CTT, score reliability is mathematically defined in terms of the proportion of true score variance to observed score variance. In this part of the chapter, we focus on *model-based* estimates of score reliability by means of CFA models. As we show below, for the one-factor model, the total amount of reliable variance provides an estimate of how precisely a certain scale score assesses a certain target construct.[3] In the one-factor model, the variance of the latent factor representing *Life Satisfaction* can be interpreted as the reliable ("construct score") variance of the scale score "Life Satisfaction." Further, *Life Satisfaction* and item-specific residual variables (i.e., e_1 to e_5) are specified to be unrelated, reflecting the idea that construct score and error score are mutually independent. Taken together, in the case of a one-factor model, the *model-based* reliability of a scale score may be defined as the proportion of variance accounted for by *one* latent target construct (e.g., *Life Satisfaction*) relative to observed score variance. In line with McDonald (1999) and Zinbarg, Yovel, Revelle, and McDonald (2006), we refer to this reliability coefficient as *omega* (ω). More formally, these ideas can be expressed as follows. When unit weights are used, a scale score X is computed by summing up p manifest scale indicators X_j: $X = X_1 + X_2 + \ldots + X_p$. When standardized model parameters are used, ω for the scale score X is computed as follows:

[3] An in-depth discussion on how to compute reliability for more complex, hierarchical constructs can be found in Brunner et al. (2012).

$$\omega = \frac{\left(\sum_{p}^{i=1} \lambda_j\right)^2}{\left(\sum_{p}^{j=1} \lambda_j\right)^2 + \sum_{p}^{j=1} Var\left(e_j\right)} \tag{2.6}$$

Here, $Var(e_j)$ is the standardized variance of the subtest-specific factor affecting the manifest variable X_j. The numerator in Eq. 2.6 represents the amount of score variance in the scale score Y that can be attributed to the variance of the factor representing the target construct. The denominator represents the total variance of the scale score, which comprises (a) the score variance accounted for by the target construct and (b) the variances attributable to the item-specific factors of the scale indicators. Values of omega can range from 0 (no reliability) to 1 (perfect reliability).

When the model parameters obtained for the one-factor model are used (Fig. 2.1), omega of the scale score "Life Satisfaction" is computed as the ratio of the variance attributable to *Life Satisfaction* to the total variance of this scale score. The total variance of the scale score "Life Satisfaction" is the sum of the variances that can be attributed to (a) *Life Satisfaction* and (b) subtest-specific factors (i.e., the sum of the variances of e_1 to e_5):

$$\omega = \left(0.78+0.63+0.75+0.63+0.62\right)^2 \Big/ \left[\begin{array}{l}\left(0.78+0.63+0.75+0.63+0.62\right)^2 \\ +\left(0.39+0.60+0.44+0.61+0.62\right)\end{array}\right] = 0.81 \tag{2.7}$$

The value of $\omega = 0.81$ represents the reliability of the scale score "Life Satisfaction" to assess individuals' levels on the latent variable *Life Satisfaction*. In other words, 81 % of the variance in this scale score is accounted for by the latent variable *Life Satisfaction* (therefore it is also called *construct reliability*). In this respect, it is important to highlight that all model-based estimates of score reliability (as well as all reliability estimates based on CTT such as Cronbach's Alpha) are population dependent. Thus, score reliability depends on how heterogeneous the sample is on the target construct (Mellenbergh, 1996).

2.3.3 Application and Interpretation of Model-Based Estimates of Reliability

Reliability is an essential test standard that should be met by virtually every observed score regardless whether the score is based on self-reports, performance on standardized tests, behavioral observations, interviews, or any other assessment method. Omega is a statistical coefficient that can be used to judge how well a set of observed measures fulfills this standard. As omega is based on parameter estimates (i.e., estimates of factor loadings and factor variances) that are derived for a certain CFA model, two vital statistical requirements need to be fulfilled: (1) Proper interpretation

of omega requires that the target model fits the empirical data well (McDonald, 1999; Yang & Green, 2010). (2) Parameter estimates need to be precise.

We first address the evaluation of model fit. In the present example, the one-factor model fits the data very well which empirically supports the interpretation of omega as a reliability coefficient for the scale score "Life Satisfaction." Notably, there has been considerable debate on which fit indices should be used and on the strategies applied to evaluate model fit (e.g., Heene et al., 2011; McDonald, 2010; West, Taylor, & Wu, 2012). Although no consensus has yet been reached, several methodologists strongly recommend comparing the preferred target model with several a priori specified and theoretically supported alternatives. This approach takes into account that cutoff values of model fit indices are model dependent, considers alternative explanations of the data, and allows some models to be ruled out while giving stronger support for others (MacCallum & Austin, 2000; West et al., 2012).

We now turn to the precision of model parameters, which is affected by two key factors. First, sample size needs to be sufficiently large to obtain trustworthy estimates of model parameters (Muthén & Muthén, 2002; Yang & Green, 2010). In general, a larger sample size is always better, and a sample size of $N \geq 200$ allows proper estimation of model parameters under a large variety of conditions (Boomsma & Hoogland, 2001). Moreover, previous simulation studies have demonstrated that trustworthy model-based reliability estimates may be obtained even with relatively small sample sizes (e.g., $N = 100$; see Zinbarg et al., 2006). It is important to note that sample size also affects the precision of the estimation of alpha (Bonett, 2003). Thus, alpha may not be preferable to omega, even in cases of small samples.

Second, parameters for CFA models are typically derived by maximum likelihood estimation, which requires continuous raw data that follow a multivariate normal distribution. However, many studies in educational and psychological research administer self-report items with a limited number of response options; therefore, the assumption that raw data are continuous may not be tenable. Moreover, empirical data frequently fail to follow a normal distribution (Micceri, 1989) and, consequently, to have a multivariate normal distribution. So what can be done? Model parameters are generally trustworthy if three conditions are fulfilled: the raw data are continuous, the sample size is reasonably large, and the assumption of multivariate normality is not severely violated. Parameter estimates are quite robust to violations of the multivariate normality assumption as long as the indicators are "reasonably" continuous. Moreover, a recent simulation study by Rhemtulla, Brosseau-Liard, and Savalei (2012) demonstrated that maximum likelihood-based estimation methods yield acceptable parameter estimates for CFA models under a wide range of conditions, even when the manifest variables contain only five response categories.

If distributional assumptions are severely violated, several routes can be taken to tackle the problem—for example, employing alternative (robust) estimation methods with less stringent distributional assumptions or transforming the input data to better match the distributional assumptions. Modern software packages used to study CFA models include robust estimation methods, such as robust maximum likelihood estimation (Satorra, 1990) and robust weighted least squares estimation

(B. O. Muthén & Kaplan, 1985). These estimation methods may yield higher precision (a) to assess model fit, (b) to compute standard errors of model parameters, and (c) in the case of robust weighted least squares, to estimate the model parameters themselves. Thus, robust weighted least squares may also be an appropriate method for analyzing item-level data from items with fewer than five response categories (Rhemtulla et al., 2012; see Wirth & Edwards, 2007 for an excellent review of factor models for item-level data). Moreover, robust maximum likelihood estimation allows the use of omega as explained in this chapter; in the case of weighted least squares estimation, score reliability may be estimated using the approaches proposed by Green and Yang (2009) or Bentler (2009, p. 142).

Alternatively, item scores that are intended to measure the same construct(s) may be integrated into parcel scores. Parcel scores may then be used as manifest measures of the latent variables in CFA models, and model parameters can be estimated by (robust) maximum likelihood procedures. Parcel scores may have several advantages over item scores: they show better distributional properties (i.e., normality), keep the ratio of observable measures to latent constructs manageable, and increase the chances of adequate model fit (Bagozzi & Edwards, 1998; Little, Rhemtulla, Gibson, & Schoemann, 2013; Marsh, Hau, Balla, & Grayson, 1998; West, Finch, & Curran, 1995). When parcel scores are applied, two key requirements need to be fulfilled: (a) The parcel scores must adequately represent the target construct(s) (Bagozzi & Edwards, 1998; Little et al., 2002). (b) The dimensional structure underlying the items needs to be taken into account. Otherwise, inaccurate parameter estimates and model fit statistics will result (Little et al., 2013).

Finally, it is interesting to note that when a one-factor model is applied (and theoretically justified, see Brunner, Nagy, & Wilhelm, 2012), omega may not only be considered as an index of score reliability but also as an index of construct validity. As noted above, omega indicates the precision with which a scale score assesses a certain target construct. This interpretation of omega converges with the concept of construct validity—the extent to which a measure assesses the construct it was designed to measure (Bollen, 1989, p. 195; McDonald, 1999, p. 63 and p. 208). Note, however, that this interpretation of omega applies only for researchers who conceive of validity as a quantitative concept and not for those who conceive of validity as a qualitative concept (i.e., a measure is or is not valid to assess a certain target construct). The latter researchers may consider two measures to be valid, but one to be more reliable (Borsboom et al., 2004, p. 1070). Omega is thus an index of reliability in terms of measurement precision only.

2.4 Challenges and Outlook

Measuring psychosocial skills is promising and has the potential to improve predictions for many important outcomes (e.g., scholastic performance, academic performance, job performance, health, etc.). However, there are still many issues that need to be investigated further to avoid some of the problems that come along with the

assessment of psychosocial skills. The main critical issues that have attracted attention during the last decades are response styles and response sets as well as the bandwidth-fidelity problem. Other interesting topics not covered here are, for example, the use of short scales (Kruyen, Emons, & Sijtsma, 2012; Kruyen, Emons, & Sijtsma, 2013a, 2013b, 2013c; but see also Gogol et al., 2014), lacking convergent validity (Pace & Brannick, 2010; Ziegler et al., 2013), or the exploration of new constructs such as the dark triad (Paulhus & Williams, 2002). The technical changes in recent years have also opened up a wide range of new assessment techniques. For example, the visual analogue scale which was not used due to reliability issues can now be applied without fearing low reliabilities (Reips & Funke, 2008). Other interesting developments are the use of videos or even virtual reality and of course web-based assessments. Unfortunately, it would be beyond the scope of this chapter to breach all of these subjects. However, despite these new developments, the general ideas outlined above should always be considered and carefully applied.

2.4.1 Response Styles and Sets

The issue of response sets and styles becomes relevant because most assessments of psychosocial skills are based on some kind of questionnaire that persons are asked to answer. The answers given to the questionnaire items on a response scale do not only reflect the underlying construct as explained above. Unfortunately, a number of response sets and styles also potentially influence the answers. Jackson and Messick (1958) differentiated between a response set and a response style (also see Paulhus, 2002).

A response style refers to a specific answering tendency a person displays regardless of item content or situation. Examples could be acquiescence, midpoint responding (MPR), or extreme point responding (EPR) (Wetzel, Böhnke, Carstensen, Ziegler, & Ostendorf, 2013; Wetzel, Carstensen, & Böhnke, 2013). Especially MPR and ERM can be dangerous in assessments. Basically, these response styles mean that two test takers with the same standing on the construct assessed but different response styles, i.e., one uses EPR and the other MPR, will receive different observed scores. This score difference does not represent an actual difference on the target construct though rendering decisions pertaining to the target construct to be false. As for acquiescence, i.e., the tendency to agree, for many years it has been suggested to use positively (e.g., I like to go out) and negatively keyed items (e.g., I like to be alone) to measure a trait (e.g., Extraversion). Supposedly, this change in direction helps preventing acquiescence. Unfortunately, it has been shown that the negatively keyed items do not only measure the intended trait but often result in method factors (Preckel, 2014) which have been related to verbal ability (e.g., Marsh, 1996). Thus, using negatively keyed items should be considered carefully.

A response set does not affect the answer regardless of content or situation. Instead, a response set occurs only for specific item contents or under specific situ-

ational demands. A classic example for a response set is faking. Ziegler, MacCann, and Roberts (2011) defined faking as:

> [A] response set aimed at providing a portrayal of the self that helps a person to achieve personal goals. Faking occurs when this response set is activated by situational demands and person characteristics to produce systematic differences in test scores that are not due to the attribute of interest. (p. 8)

In statistical terms, the axiom of classical test theory could be expanded to X = T + E + I(P,S) with I(P,S) meaning an interaction between person and situational demand (Heggestad, George, & Reeve, 2006; Ziegler & Bühner, 2009). Thus, faking only occurs in specific situations (Ellingson, 2011). However, research shows that faking does affect reliability (MacCann, 2013) and validity (MacCann, Ziegler, & Roberts, 2011).

Methods that help to identify and model response sets and response styles are, for example, mixed Rasch models (Rost, Carstensen, & Von Davier, 1997) and factor mixture models (Lubke & Muthén, 2005). In both cases the underlying idea is that answers to an item are caused by a latent construct plus a specific response set or style. In order to model both, mixed Rasch models and factor mixture models combine a method to model a construct (i.e., Rasch model or structural equation modeling, respectively) with a latent class analysis. Thus, the construct is modeled as a continuous latent variable and response set and styles as a class membership. Both methods have successfully been applied in the context of response styles (Wetzel, Böhnke et al., 2013; Ziegler & Kemper, 2013) and response sets (Zickar, Gibby, & Robie, 2004). Potentially, these approaches could also help to correct scores for response sets and styles. However, this approach still is in its fledgling stages (Zickar & Sliter, 2011).

Some new or newly discovered item types or scoring techniques like multidimensional pairwise preference items and the use of anchoring vignettes are a more promising road to deal at least with response sets like faking. *Multidimensional pairwise preference items* often are quadruples, i.e., four different statements are presented in one item, and the test taker has to choose which of the statements is most like her or him and which one is least like her or him. The statements refer to different constructs (multidimensional) and are usually matched in terms of their desirability. Using generalized graded unfolding models (Stark, Chernyshenko, & Drasgow, 2011) or Rasch models (Brown & Maydeu-Olivares, 2011), scores can be derived that can be used to compare individuals with each other.

Anchoring vignettes (King, Murray, Salomon, & Tandon, 2004) rescale test taker answers based on their answers to standard vignettes. Thus, it is observed how a person uses the rating scale in general. This information is then used to transform the answers and thereby correct for response styles and sets. The technique has successfully been applied in educational settings (Kyllonen & Bertling, 2013) and especially in cross-country studies (Bolt, Lu, & Kim, 2014; Mõttus et al., 2012).

Thus, response sets and styles have been identified as a threat to the assessment of psychosocial skills. Recent research has not only allowed to model their impact but also promises to find methods to deal with them. Nevertheless, future research will have to intensify the focus on these issues.

2.4.2 *Bandwidth-Fidelity Problem and Test-Criterion Correlations*

The quality of predicting actual behavior from using psychosocial skills or any other individual characteristic depends on many factors. One of the factors has become known as the bandwidth-fidelity problem (Hogan & Roberts, 1996; Ones & Viswesvaran, 1996). Generally, this means that a score derived from a less abstract measure (see also level of symmetry above), for example, a personality facet score, often is a better predictor of behavior than the more abstract domain score (Paunonen & Ashton, 2001; Ziegler, Danay, Schölmerich, & Bühner, 2010). However, the less abstract score is often based on fewer items resulting in lower reliability estimates. Still, considering the test-criterion correlations, the focus on less abstract facets seems to be a promising research track. It will be vital, though not to lose track of the overarching nomological network within which all the facets lie. This means lower-order facets should always be integrated into existing higher-order structures. Thus, it should be defined which higher-order constructs the facets belong to or how the facets relate to known higher-order structures like the Big Five. Otherwise, countless new constructs are likely to emerge, and chances are that some of them are redundant to a large extent.

Another issue relevant for test-criterion correlation has attracted less research so far. Most people would agree that their behavior is influenced by the situation. In most cultures, even an extraverted person will not act jovially and joyfully at a funeral because the situation requires a deferential and quiet behavior. This interactionism is an accepted phenomenon (Bowers, 1973; Mischel, 2004). Thus, personality traits do not always manifest themselves in a given situation. Tett and Burnett (2003) have proposed a theory which explains this phenomenon in the work context. Their trait activation theory states that there are specific situational features (e.g., demands, distracters, constraints, releaser, and facilitators) that enable or block the manifestation of a trait and thereby influence test-criterion correlations. Consequently the same trait score can be a valid predictor in one situation (or job) but not in another (Ziegler et al., 2014). This kind of research shows that in order to substantially improve the quality of predictions the incorporation of situation perception is necessary (Rauthmann, 2012).

2.5 Conclusion

Summing up, the present chapter gave a short overview of important test standards and their implications for selecting and constructing psychological measures. Moreover, we introduced basic and more advanced psychometric models and their relevance for estimating score reliability. Finally, we outlined challenges for psychosocial assessment and described some potential remedies. Obviously, each of the topics has much more important details to offer than would be possible to present

within a single chapter. Thus, we would like to advise the reader to engage with the literature cited here and related literature not cited here. Only a thorough understanding of general psychometric principles as well as a strict adherence to test standards allows a responsible approach to the assessment of psychosocial skills.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association, American Psychological Association, & National Council on Measurement in Education.

Bagozzi, R. P., & Edwards, J. R. (1998). A general approach for representing constructs in organizational research. *Organizational Research Methods, 1*, 45–87.

Barrick, M. R., & Mount, M. K. (1991). The BIG five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1–26.

Bentler, P. M. (2009). Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika, 74*, 137–143.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Bolt, D. M., Lu, Y., & Kim, J.-S. (2014). Measurement and control of response styles using anchoring vignettes: A model-based approach. *Psychological Methods*. doi:10.1037/met0000016.

Bonett, D. G. (2003). Sample size requirements for testing and estimating coefficient alpha. *Journal of Educational and Behavioral Statistics, 27*, 335–340.

Boomsma, A., & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation models: Present and future. A Festschrift in honor of Karl Jöreskog* (pp. 139–168). Lincolnwood, IL: Scientific Software International.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review, 111*, 1061–1071.

Bowers, K. S. (1973). Situationism in psychology—Analysis and a critique. *Psychological Review, 80*(5), 307–336.

Brogden, H. E., & Taylor, E. K. (1950). The theory and classification of criterion bias. *Educational and Psychological Measurement, 10*(2), 159–186.

Brown, A., & Maydeu-Olivares, A. (2011). Item response modeling of forced-choice questionnaires. *Educational and Psychological Measurement, 71*(3), 460–502.

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality, 80*, 796–846. doi:10.1111/j.1467-6494.2011.00749.x.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62*, 193–217.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81–105.

Cattell, R. B. (1958). What is“ objective” in“ objective personality tests?”. *Journal of Counseling Psychology, 5*, 285–289.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98–104.

Cronbach, L. J. (1947). Test “reliability”: Its meaning and determination. *Psychometrika, 12*(1), 1–16.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297–334.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*, 71–75.

Eid, M., Nussbeck, F., Geiser, C., Cole, D., Gollwitzer, M., & Lischetzke, T. (2008). Structural equation modeling of multitrait-multimethod data: Different models for different types of methods. *Psychological Methods, 13*(3), 230–253.

Ellingson, J. E. (2011). People fake only when they need to fake. In M. Ziegler, C. MacCann, & R. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 19–33). New York: Oxford University Press.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory*. New York: Psychology Press.

Fan, X. (2003). Two approaches for correcting correlation attenuation caused by measurement error: Implications for research practice. *Educational and Psychological Measurement, 63*, 915–930.

Gogol, K., Brunner, M., Goetz, T., Martin, R., Ugen, S., Keller, U., et al. (2014). "My Questionnaire is Too Long!" The assessments of motivational-affective constructs with three-item and single-item measures. *Contemporary Educational Psychology, 39*(3), 188–205.

Green, S. B., & Yang, Y. (2009). Reliability of summed item scores using structural equation modeling: An alternative to coefficient alpha. *Psychometrika, 74*, 155–167.

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods, 6*, 430–450.

Heene, M., Hilbert, S., Draxler, C., Ziegler, M., & Bühner, M. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: A cautionary note on the usefulness of cutoff values of fit indices. *Psychological Methods, 16*, 319–336. doi:10.1037/a0024917.

Heene, M., Hilbert, S., Freudenthaler, H. H., & Bühner, M. (2012). Sensitivity of SEM fit indexes with respect to violations of uncorrelated errors. *Structural Equation Modeling: A Multidisciplinary Journal, 19*, 36–50.

Heggestad, E. D., George, E., & Reeve, C. L. (2006). Transient error in personality scores: Considering honest and faked responses. *Personality and Individual Differences, 40*, 1201–1211. doi:10.1016/j.paid.2005.10.014.

Hogan, J., & Roberts, B. (1996). Issues and non-issues in the fidelity-bandwidth trade-off. *Journal of Organizational Behavior, 17*(6), 627–637.

Hoyle, R. H., & Smith, G. T. (1994). Formulating clinical research hypotheses as structural equation models: A conceptual overview. *Journal of Consulting and Clinical Psychology, 62*, 429–440.

Jackson, D. N., & Messick, S. (1958). Content and style in personality-assessment. *Psychological Bulletin, 55*(4), 243–252.

King, G., Murray, C. J., Salomon, J. A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *The American Political Science Review, 98*(1), 191–207.

Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology, 50*, 537–567.

Kruyen, P. M., Emons, W. H., & Sijtsma, K. (2012). Test length and decision quality in personnel selection: When is short too short? *International Journal of Testing, 12*(4), 321–344.

Kruyen, P. M., Emons, W. H. M., & Sijtsma, K. (2013a). Assessing individual change using short tests and questionnaires. *Applied Psychological Measurement, 38*(3), 201–216.

Kruyen, P. M., Emons, W. H. M., & Sijtsma, K. (2013b). Shortening the S-STAI: Consequences for research and clinical practice. *Journal of Psychosomatic Research, 75*(2), 167–172.

Kruyen, P. M., Emons, W. H., & Sijtsma, K. (2013c). On the shortcomings of shortened tests: A literature review. *International Journal of Testing, 13*(3), 223–248.

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika, 2*(3), 151–160.

Kyllonen, P. C., & Bertling, J. P. (2013). Innovative questionnaire assessment methods to increase cross-country comparability. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (pp. 277–285). Boca Raton: CRC Press.

Lindqvist, E., & Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment. *American Economic Journal: Applied Economics, 3*, 101–128. doi:10.1257/app.3.1.101.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*(2), 151–173.

Little, T. D., Rhemtulla, M., Gibson, K., & Schoemann, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods, 18*(3), 285–300. doi:10.1037/a0033266.

Loevinger, J. (1954). The attenuation paradox in test theory. *Psychological Bulletin, 51*, 493.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Lubke, G. H., & Muthén, B. O. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods, 10*(1), 21–39.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51*, 201–226.

MacCann, C. (2013). Instructed faking of the HEXACO reduces facet reliability and involves more Gc than Gf. *Personality and Individual Differences, 55*(7), 828–833.

MacCann, C., Ziegler, M., & Roberts, R. (2011). Faking in personality assessment: Reflections and recommendations. In M. Ziegler, C. MacCann, & R. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 309–329). New York: Oxford University Press.

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology, 79*, 280.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology, 70*(4), 810–819.

Marsh, H. W., Hau, K. T., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research, 33*, 181–220.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577–605.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.

McDonald, R. P. (2010). Structural models and the art of approximation. *Perspectives on Psychological Science, 5*, 675–686.

Mellenbergh, G. J. (1996). Measurement precision in test score and item response models. *Psychological Methods, 1*, 293–299.

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin, 105*(1), 156–166.

Miller, J. D., Gaughan, E. T., Maples, J., & Price, J. (2011). A comparison of agreeableness scores from the Big Five inventory and the NEO PI-R: Consequences for the study of narcissism and psychopathy. *Assessment, 18*(3), 335–339.

Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology, 55*, 1–22.

Mõttus, R., Allik, J., Realo, A., Rossier, J., Zecca, G., Ah-Kion, J., et al. (2012). The effect of response style on self-reported conscientiousness across 20 countries. *Personality and Social Psychology Bulletin, 38*(11), 1423–1436.

Mussel, P. (2010). Epistemic curiosity and related constructs: Lacking evidence of discriminant validity. *Personality and Individual Differences, 49*(5), 506–510.

Muthén, B. O., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology, 38*, 171–189.

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling, 9*, 599–620.

Ones, D., & Viswesvaran, C. (1996). Bandwidth-fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior, 17*(6), 609–626.

Pace, V. L., & Brannick, M. T. (2010). How similar are personality scales of the "same" construct? A meta-analytic investigation. *Personality and Individual Differences, 49*(7), 669–676.

Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Mahwah, NJ: Lawrence Erlbaum Associates.

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality, 36*, 556–563.

Paunonen, S. V., & Ashton, M. C. (2001). Big five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology, 81*(3), 524–539.

Preckel, F. (2014). Assessing need for cognition in early adolescence: Validation of a German adaption of the Cacioppo/Petty scale. *European Journal of Psychological Assessment, 30*, 65–72. doi:10.1027/1015-5759/a000170.

Rajaratnam, N., Cronbach, L., & Gleser, G. (1965). Generalizability of stratified-parallel tests. *Psychometrika, 30*(1), 39–56.

Rauthmann, J. F. (2012). You say the party is dull, I say it is lively a componential approach to how situations are perceived to disentangle perceiver, situation, and perceiver × situation variance. *Social Psychological and Personality Science, 3*(5), 519–528.

Reips, U., & Funke, F. (2008). Interval-level measurement with visual analogue scales in internet-based research: VAS generator. *Behavior Research Methods, 40*, 699–704.

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods, 17*(3), 354–373. doi:10.1037/a0029315.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2*, 313–345. doi:10.1111/j.1745-6916.2007.00047.x.

Rost, J., Carstensen, C. H., & Von Davier, M. (1997). Applying the mixed Rasch model to personality questionnaires. In J. In Rost & R. E. Langeheine (Eds.), *Applications of latent trait and latent class models in the social sciences*. New York: Waxmann.

Satorra, A. (1990). Robustness issues in structural equation modeling: a review of recent developments. *Quality and Quantity, 24*, 367–386.

Schumacher, J., Klaiberg, A., & Brähler, E. (Eds.). (2003). *Diagnostische Verfahren zu Lebensqualität und Wohlbefinden* [Diagnostic methods for assessing quality of life and subjective well-being]. Göttingen: Hogrefe.

Shavelson, R. J., & Webb, N. M. (2006). Generalizability theory. In J. L. Green, G. Camilli, & P. B. Elmore (Eds.), *Handbook of complementary methods in education research*. (pp. 309–322). Mahwah, NJ: Lawrence Erlbaum Associates.

Slaney, K. L., & Maraun, M. D. (2008). A proposed framework for conducting data-based test analysis. *Psychological Methods, 13*, 376–390.

Spengler, M., Lüdtke, O., Martin, R., & Brunner, M. (2013). Personality is related to educational outcomes in late adolescence: Evidence from two large-scale achievement studies. *Journal of Research in Personality, 47*(5), 613–625. doi:10.1016/j.jrp.2013.05.008.

Stark, S., Chernyshenko, O. S., & Drasgow, F. (2006). Detecting differential item functioning with confirmatory factor analysis and item response theory: Toward a unified strategy. *Journal of Applied Psychology, 91*, 1292.

Stark, S., Chernyshenko, O., & Drasgow, F. (2011). Constructing fake-resistant personality tests using item response theory: High-stakes personality testing with multidimensional pairwise preferences. In M. Ziegler, C. MacCann, & R. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 214–239). New York: Oxford University Press.

Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology, 88*(3), 500–517.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4–69.

Walker, C. M. (2011). What's the DIF? Why differential item functioning analyses are an important part of instrument development and validation. *Journal of Psychoeducational Assessment, 29*, 364–376. doi:10.1177/0734282911406666.

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56–75). Thousand Oaks, CA: Sage.

West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). New York: Guilford Press.

Wetzel, E., Böhnke, J. R., Carstensen, C. H., Ziegler, M., & Ostendorf, F. (2013). Do individual response styles matter? *Journal of Individual Differences, 34*(2), 69–81.

Wetzel, E., Carstensen, C. H., & Böhnke, J. R. (2013). Consistency of extreme response style and non-extreme response style across traits. *Journal of Research in Personality, 47*(2), 178–189.

Wirth, R. J., & Edwards, M. C. (2007). Item factor analysis: Current approaches and future directions. *Psychological Methods, 12*, 58–79.

Yang, Y., & Green, S. B. (2010). A note on structural equation modeling estimates of reliability. *Structural Equation Modeling, 17*, 66–81.

Zickar, M. J., Gibby, R. E., & Robie, C. (2004). Uncovering faking samples in applicant, incumbent, and experimental data sets: An application of mixed-model item response theory. *Organizational Research Methods, 7*(2), 168–190.

Zickar, M. J., & Sliter, K. A. (2011). Searching for unicorns: Item response theory-based solutions to the faking problem. In M. Ziegler, C. MacCann, & R. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 113–130). New York: Oxford University Press.

Ziegler, M., Bensch, D., Maaß, U., Schult, V., Vogel, M., & Bühner, M. (2014). Big Five facets as predictor of job training performance: The role of specific job demands. *Learning and Individual Differences, 29*(1), 1–7.

Ziegler, M., Booth, T., & Bensch, D. (2013). Getting entangled in the nomological net. *European Journal of Psychological Assessment, 29*(3), 157–161.

Ziegler, M., & Bühner, M. (2009). Modeling socially desirable responding and its effects. *Educational and Psychological Measurement, 69*(4), 548–565.

Ziegler, M., Danay, E., Schölmerich, F., & Bühner, M. (2010). Predicting academic success with the Big 5 rated from different points of view: Self-rated, other rated and faked. *European Journal of Personality, 24*(4), 341–355.

Ziegler, M., & Kemper, C. (2013). Extreme response style and faking: Two sides of the same coin? In P. Winker, R. Porst, & N. Menold (Eds.), *Interviewers' deviations in surveys: Impact, reasons, detection and prevention (Schriften Zur Empirischen Wirtschaftsforschung)* (pp. 221–237). Frankfurt a. M: Peter Lang Gmbh.

Ziegler, M., MacCann, C., & Roberts, R. R. (2011). Faking: Knowns, unknowns, and points of contention. In M. Ziegler, C. MacCann, & R. R. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 3–16). New York: Oxford University Press.

Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for ωh. *Applied Psychological Measurement, 30*, 121–144.

Chapter 3
Aligning Mission and Measurement

Steven E. Stemler and Mary DePascale

3.1 Introduction

There is a deep disconnect between what K-12 schools in the USA articulate as their primary mission and how those schools are held accountable for their performance. Specifically, principals, teachers, parents, employers, and policy makers all believe that schools should be doing more than just teaching students to read, write, and do math (Stemler, Bebell, & Sonnabend, 2011). There is a strong consensus that students should also develop emotional skills and learn the skills associated with effective citizenship (Stemler & Bebell, 2012). These broader skills are viewed as fundamental to the core mission of schools and, in some cases, hold an even higher priority than the basic cognitive elements. Yet, the current push for accountability coming from the federal level, in the form of standardized testing (most recently the Common Core State testing), relates exclusively to cognitively oriented, domain-specific knowledge (Common Core State Standards Initiative, 2010a, 2010b).

The most common arguments we have encountered for dismissing the idea of assessing broader skills are as follows: (1) they are nice "add-ons" to a student's education, but are not fundamental to it; (2) these skills cannot be assessed because they are too subjective. Our primary goal in this chapter will be to address these two arguments directly.

In order to ensure that measurement is aligned with mission, we must first systematically examine the mission of schooling. Therefore, this chapter begins with a review of the literature related to school purpose. We draw on various sources of data (e.g., surveys, legal documents, school mission statements) from a diverse

S.E. Stemler (✉) • M. DePascale
Department of Psychology, Wesleyan University,
207 High Street, Middletown, CT 06437, USA
e-mail: steven.stemler@wesleyan.edu; mdepascale@wesleyan.edu

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_3

range of constituents (e.g., courts, businesses, schools) in order to examine empirically the core competencies these constituencies expect schools to develop in students.

In the second part of the chapter, our main objective is to illustrate for the reader how different noncognitive competencies of interest have been measured in the educational and psychological literature. While many instruments that purport to measure these important noncognitive skills have been developed, those included in this chapter have demonstrated strong psychometric evidence, are typically aimed at K-12 students, and represent a diversity of approaches to measuring the core competency. We comment on the nature of the instruments (e.g., self-report vs. performance), provide sample items, highlight information related to their reliability and validity, and refer the reader to where, specifically, the complete instruments may be found. Our review is not intended to be exhaustive; rather, the main point we wish to convey is that a wide variety of psychometrically sound measures currently exist for nearly any competency one wishes to measure. In addition, new measures are being introduced into the literature constantly.

We conclude this chapter by proposing a new approach to accountability that aligns mission and measurement. We argue that because not all schools aim to achieve the same outcomes, an ideal accountability system would be one in which schools are held accountable for those objectives they aim to achieve. We outline one model for what that could look like within the current political context.

3.2 The Mission of Schools

The purpose of schooling is a topic that has been debated by philosophers, politicians, academics, legal courts, businesses, parents, and students since the inception of formal schooling. Philosophers as diverse as Aristotle, Emerson, Plato, Locke, Confucius, Dewey, Counts, and Adler have written about this topic (Noddings, 1995; Reed & Johnson, 1996), and it is a favorite topic of sociologists (deMarrais & LeCompte, 1995; Labaree, 1997) and historians alike (Goodlad, 1979; Tyack, 1988). We will now briefly review some of these perspectives.

Legal Perspectives Despite the growing federal influence in American education over the past 30 years, which has largely emphasized mainly cognitive development and vocational preparation (e.g., A Nation at Risk, 1983; Goals 2000; NCLB, 2001; Common Core State Standards Initiative, 2010a, 2010b), court rulings have largely maintained individual states' rights to self-determine the goals of their public educational systems.

Within the past 25 years, courts in states ranging from Kentucky to Massachusetts have outlined what they believe to be the purposes of schooling from a legal perspective. In 1989 (Rose v. Council for Better Education, 1989), the Kentucky State Supreme Court ordered the General Assembly to reform the property tax system and provide funding "sufficient to provide each child in Kentucky an adequate education." In defining an adequate education, the court enumerated seven learning

goals that have been widely cited as precedent and subsequently adopted by other states (e.g., McDuffy v. Secretary, 1993). The seven distinct components of education outlined by the court include the development of (i) sufficient oral and written communication skills to enable a student to function in a complex and readily changing civilization; (ii) sufficient knowledge of economic, social, and political systems to enable students to make informed choices; (iii) sufficient understanding of government processes to enable the student to understand the issues that affect his or her community, state, and nation; (iv) sufficient self-knowledge and knowledge of his or her mental and physical wellness; (v) sufficient grounding in the arts to enable each student to appreciate his or her cultural and historical heritage; (vi) sufficient training or preparation for advanced training in either academic or vocational fields so as to enable each child to choose and pursue life work intelligently; and (vii) sufficient level of academic or vocational skills to enable public school students to compete favorably with their counterparts in surrounding states, in academics, or in the job market.

In recognizing the many goals of public education, the Kentucky precedent emphasizes that public schooling should not simply focus on cognitive outcomes. Specifically, the courts in Kentucky, Massachusetts, and many other states articulate an equal emphasis on a variety of student outcomes including cognitive, civic, and emotional development. As a practical consequence, states have slowly begun to require exit exams in broader domains, such as citizenship, US history, or social studies (Kemler, 2015; McIntosh, 2012), as well as mathematics, science, reading, or writing.

Business Perspectives The perspectives of citizens (Immerwahl, 2000) as well as businesses have been studied via the use of large-scale surveys. One recent large-scale study of employers conducted by the Association of American Colleges and Universities (2010) revealed that the educational outcomes considered important by employers include the ability to communicate effectively, orally, and in writing (89 % of employers surveyed); critical thinking and analytical reasoning skills (81 %); the ability to apply knowledge and skills to real-world settings through internships or other hands-on experiences (79 %); the ability to connect choices and actions to ethical decisions (75 %); the ability to analyze and solve complex problems (75 %); teamwork skills and the ability to collaborate with others in a diverse group setting (71 %); the ability to innovate and be creative (70 %); the ability to locate, organize, and evaluate information from multiple sources (68 %); the ability to work with numbers and understand statistics (63 %); an understanding of the role of the USA in the world (57 %); an appreciation for cultural diversity in America and other countries (57 %); and civic knowledge, civic participation, and community engagement (52 %).

Similarly, an evaluation of Recruiting Trends (Gardner, 2007), a publication based on information supplied by hundreds of companies and organizations concerning the recruitment of recent college graduates, reveals what specific skills employers are seeking in their recruits. In 2002–2003, ethics and integrity were considered the most important competencies. The following year, employers expressed their preference for college students to have better developed skills in communication, personal attributes (work ethic, flexibility, initiative, and motivation),

teamwork, interpersonal skills, and learning (willing to learn continuously new skills and ideas). The next year, in the 2005–2006 issue, employers placed a high value on geographic awareness and a global understanding of events as they pertain to the company and industrial sector. Likewise, Casner-Lotto and Barrington (2006) surveyed 400 employers across the USA and found that the most important skills they felt were needed to succeed in the workplace included (1) professionalism/work ethic, (2) oral and written communication, (3) teamwork/collaboration, and (4) critical thinking/problem solving.

School Perspectives The study of school mission statements has emerged as a productive approach to empirically studying what schools themselves articulate as their core objectives (Bebell & Stemler, 2002; Schmitt, 2012; Stemler & Bebell, 1999, 2012; Stemler et al., 2011; Stober, 1997). A wide range of school effectiveness research has consistently shown that commitment to a shared mission statement is one of the leading factors differentiating more effective schools from less effective schools (Kurland, Peretz, & Hertz-Lazarowtiz, 2010; Renchler, 1991; Renihan, Renihan, & Waldron (1986); Rutter & Maughan, 2002; Teddlie & Reynolds, 2000). Although we would not argue that the mission is the only indicator of a school's cultural values, we do argue that it provides a straightforward and accessible indicator.

School mission statements represent a useful source of data for gaining access to the school perspective on matters of purpose for at least four reasons. First, nearly all major school accrediting bodies require a mission statement from schools seeking accreditation (AdvanceEd, 2006). Indeed, the very first standard articulated by the nation's largest secondary school accreditation body requires that:

> Schools establish and communicate a shared purpose and direction for improving the performance of students and the effectiveness of the school. In fulfillment of the standard, the school: (i) establishes a vision for the school in collaboration with its key stakeholders, (ii) communicates the vision and purpose to build stakeholder understanding and support, (iii) identifies goals to advance the vision, (iv) ensures that the school's vision and purpose guide the learning process, and (v) reviews its vision and purpose systematically and revises them when appropriate. (AdvanceEd, 2006, p. 9)

As recognized by most accrediting boards (as well as many business, civic, and private organizations in America), mission statements represent an important summation or distillation of an organizatio's core goals represented by concise and simple statements that communicate broad themes. Furthermore, school mission statements are one of the only written documents outlining purpose that nearly all schools have. The fact that nearly all American schools have a mission statement thus provides a common measure allowing for systematic comparison across diverse institutions. Second, school mission statements tend to be publicly available and easily accessible, making them well suited for study, particularly in the age of online data collection. Third, research has demonstrated that mission statements can be systematically and reliably coded by applying content analysis techniques (Bebell & Stemler, 2004; Berleur & Harvanek, 1997; Schmitt, 2012; Stemler & Bebell, 1999; Stober, 1997; Stemler et al., 2011). In 1999, Stemler and Bebell

introduced a coding rubric for school mission statements, using emergent analytic coding, that allowed mission statements to be classified into major thematic categories (e.g., social development, cognitive development, emotional development, civic development, physical development) as well as 33 specific subcategories. Across a series of studies, these authors found their rating system exhibited median consensus estimates of interrater reliability ranging from 0.77 to 0.80 (Bebell & Stemler, 2004; Stemler & Bebell, 1999). Thus, a given school mission statement can be dichotomously coded on 11 independent traits, which then allows for quantitative/statistical analyses of these traits across samples of schools.

In 2012, Stemler and Bebell studied the mission statements of a wide variety of school types in the USA. Schools were drawn from public and private K-12 education, Montessori schools, charter schools, Waldorf schools, Native American schools, and vocational schools. Even across this wide variety of school types, the data suggested a general convergence on the importance of cognitive, emotional, and civic goals; however, there were notable differences by school type. For example, none of the Montessori schools sampled included any mention of any competencies related to citizenship. Further, the Waldorf schools tended to emphasize broad, emotionally based competencies and made little explicit mention of cognitive outcomes. Vocational schools tended to focus not only on job preparation but also on cognitive outcomes and spoke little of emotional or civic elements. Figure 3.1 provides four sample mission statements drawn from different school types. Note the emphasis on individual cognitive and emotional competencies in the Montessori example (and the absence of reference to the civic), the emphasis on a wide range of emotionally oriented competencies in the Waldorf school, the nearly exclusive focus on cognitive outcomes from the charter school, and the civic emphasis of the public high school.

International Perspectives Although large-scale internationally comparative studies of achievement began in the 1970s (Stemler, 2001), the policy discussions invoking the results of large-scale tests of achievement reached a fever pitch in the 1990s. Since that time, there has been growing public and political interest in comparative educational studies (e.g., Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading Literacy Study (PIRLS), Program for International Student Assessment (PISA)). Although these international comparative studies historically have focused on traditional cognitive outcomes such as mathematics, science, and writing, the international community has begun to recognize the importance of broader competencies and have recently begun to incorporate them into their assessment battery. PISA 2018, for example, will require an assessment of teamwork skills, self-efficacy, and cultural competence in addition to their more traditional core areas (see Chap. 14).

Summary A wide variety of empirical data has been collected from a variety of constituents, including schools, parents, businesses, and courts. The data comes in a variety of forms (e.g., surveys, legal documents, mission statements), but it is remarkably consistent in pointing to the same conclusion. Constituents believe that

Chesterfield Montessori School; Chesterfield, Missouri
Chesterfield Montessori School offers an authentic Montessori education that honors children's individuality. Our peaceful environment and compassionate staff nurtures respect for self and others, fosters a strong sense of community, and stimulates independent thinking. Students carry with them a solid record of academic achievement, a belief in the dignity of work, and a sense of responsibility for their own development as happy and productive human beings.

The Bay School (Waldorf School); Blue Hill, Maine
The Bay School's mission is to provide an education that engages and nurtures the whole child, inspiring a balanced growth of heart, mind, body, and spirit. We are committed to developing in our students inner confidence, responsibility, self-motivation, a love of learning, imagination, creativity, and intellectual clarity. The educational ideals and values of the school, rooted in the Waldorf tradition, create a community of children, alumni, parents, and faculty imbued with reverence for others and the natural world.

*Jumoke Academy Charter School;*Hartford, Connecticut
The mission of Jumoke Academy Charter School is to prepare children to successfully compete in the global marketplace despite the social and economic challenges they may presently face. The academy is dedicated to rigorous academic and social standards achieved by holding high expectations for all students during challenging instruction.

The concept of "Jumoke" is central to the academy's mission to provide a safe and nurturing environment for its children while providing high quality instruction. Students in PreK-8th grade will be offered a developmentally appropriate curriculum and an enriched program of extended day activities which addresses the unique talents and background of each child in the areas of science, mathematics, language arts, technology, physical education, music and art enrichment.

Chapin High School (Recognized by the US Dept of Ed as a Blue Ribbon Public High School); Chapin, South Carolina
The mission of Chapin High School of Lexington Richland School District Five, in partnership with the community, is to provide challenging curricula with high expectations for learning that develop productive citizens who can solve problems and contribute to a global society.

Fig. 3.1 Example of school mission statements

schools exist to develop a wide variety of cognitive and noncognitive competencies. Furthermore, there is no evidence to suggest that cognitive outcomes are of greater importance to these groups or that noncognitive competencies are somehow secondary or "add-ons" to the core educational experience. To the contrary, if precedence exists, it seems to favor the development of noncognitive competencies. If schools value these skills, then should not these skills also be assessed to determine whether or not students are developing them as a result of their schooling? We now turn to the question of how to measure these important broader skills and competencies.

3.3 Measuring Broader Skills and Competencies

One reason that cognitive skills may be perceived as taking precedence over broader, noncognitive measures is because the measurement of cognitive skills has a long history in the literature, dating back over 100 years with the first standardized IQ tests (Birney & Stemler, 2007; Ciancialo & Sternberg, 2004). In addition, it is cognitive skills, and those skills alone, that are emphasized within the current high-stakes testing policies in the USA associated with federal educational policies such as Race to the Top (http://www2.ed.gov/programs/racetothetop/index.html) and the Common Core State Standards (2010a, 2010b). The reason that cognitive skills such as critical thinking are given precedence in federal policies is likely attributable to the perception that noncognitive skills are subjective and cannot be assessed as reliably as cognitive skills such as critical thinking. However, there has been substantial research in the past half century devoted to the measurement of civic, emotional, and personal skills (see Chap. 2) The argument that there is no way to measure these broader competencies is simply a canard.

The remainder of this chapter will be devoted to debunking this fallacy by citing assessment efforts associated with some of the core skills and competencies schools say they aim to develop. Given the large number of potential competencies that could be included and the limitations of space associated with this chapter, we have chosen to focus on the two core elements that Stemler and Bebell (Stemler et al., 2011; Stemler & Bebell, 2012) have shown to have equal status to the cognitive domain: those of emotional development and citizenship. Note that each of these domains represents a broad category that itself consists of many other competencies. Each of the various competencies has a substantial literature associated with it, and we fully recognize that our efforts here are merely scratching the surface. Furthermore, specialists in these areas may well dispute our grouping of the competencies as falling under the "Emotional" or "Civic" umbrella. We admit that these are rough categorizations; however, we will generally follow the rubric set forth by Stemler et al. (2011) which specifies subdomains associated with each broad category. This rubric has been demonstrated to have strong interrater reliability across a variety of studies. The instruments we have chosen to include in our review were purposefully selected to meet two or more of the following three criteria: (1) they possess strong psychometric properties; (2) they exemplify a variety of different assessment techniques (e.g., self-report, ability test, observations); and (3) they are relevant to a K-12 audience. The instruments we highlight are not the only instruments that potentially meet these criteria, but they do suffice to illustrate our broader point that objective, quantitative, and psychometrically sound measures exist for nearly any competency one wishes to consider.

3.4 Emotional Development

Emotional development was the second most frequently cited theme to emerge from Stemler et al.'s (2011) analysis of a true random sample of 500 US high school mission statements across ten geographically and politically diverse states. Fully 55 %

of the 421 randomly selected public high schools mentioned emotional development, as compared to 58 % who mentioned citizenship and 53 % who mentioned cognitive development. Further, in a subsequent study, Stemler and Bebell (2012) found that across a nonrandom sample of 111 schools from ten different school types (e.g., public, parochial, Montessori, charter, Waldorf, Native American), emotional development was the most frequently cited theme (66 %). Consequently, we being with a review of competencies associated with the domain of emotional development. According to the rubric outlined in the studies by Stemler and colleagues, these include (1) emotional intelligence; (2) empathy; (3) self-esteem, self-confidence, and self-efficacy; (4) motivation; and (5) self-directed learning. At the end of this section, we present a table that summarizes the competency under investigation, the name of the instruments used to assess the competency, example items, their psychometric properties, and information on how to obtain the instrument.

Emotional Intelligence There are two main approaches to measuring emotional intelligence that pervade the educational and psychological literature (Zeidner, Matthews, & Roberts, 2009; see also Chap. 11). One approach is based on self-report assessments of emotional intelligence and follows in the tradition of personality assessment. Perhaps the most popular instrument in this tradition is the Bar-On EQ-i. The instrument contains 133 items, each of which is rated on a 5-point Likert scale ("very seldom or not true of me," "seldom true of me," "sometimes true of me," "often true of me," "very often true of me or true of me"). Prompts ask questions similar to the following: "I can tell when someone is angry," "I know how to get someone into a good mood," and "When I need to get a job done, I can get myself into the right emotional state to get it done." This instrument has been used widely and normed on a large database of participants. The instrument has shown solid levels of internal consistency reliability, ranging from 0.69 to 0.86, and test–retest reliability coefficients ranging from 0.55 to 0.82 across the published technical manuals (Bar-on, 1997). There is a youth version of the EQ-i that is suitable for use with elementary, middle, and high school students. It has been normed against children in North America, and norms are provided separately for boys and girls across four age groups. The EQ-i youth form has been shown to be a significant predictor of academic achievement (Parker et al., 2004)

A second major approach to the measurement of emotional intelligence is through the use of performance-based assessments. In this domain, the Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT) is the most prominent test in use. In contrast to the Bar-On, this test consists of four dimensions: (1) perceiving emotions, (2) using emotions to facilitate thought, (3) understanding emotions, and (4) managing emotions. Rather than being a self-report measure, the MSCEIT is viewed as an ability test, in which there are correct and incorrect responses to each of the items. The items vary in type of presentation. In one subtest, participants are shown a face and asked to rate the extent to which different emotions are present or absent from the facial expression. In another test, individuals are presented with a particular scenario (e.g., meeting the parents of your new love interest for the first time) and asked what emotions would be most useful in that context. The instru-

ment has shown very high levels of internal consistency reliability (alpha = 0.91 full scale) as well as strong test–retest reliability (alpha = 0.86 full scale) in the published technical manuals (Mayer, Salovey, & Caruso, 2002).

Other tests measure components of emotional intelligence, such as emotional management. MacCann, Fogarty, Zeidner, and Roberts (2011) have developed a Situational Test of Emotional Management for Youth (STEM-Y) aimed at assessing this particular element of emotional intelligence. An example item from the test is: "You and James sometimes help each other with homework. After you help James on a difficult project, the teacher is very critical of this work. James blames you for his bad grade. You respond that James should be grateful, because you were doing him a favor. What would you do in this situation? (a) Tell him from now on he has to do his own homework. (b) Apologize to him. (c) Tell him 'I am happy to help, but you are responsible for what you turn in.' (d) Don't talk to him." They administered the test to 383 eighth grade students and found significant positive associations with GPA and problem-focused coping.

In addition, the Trait Emotional Intelligence Questionnaire (TEIQue) is often used as a measure of emotional intelligence. It is available in both a child form (TEIQue-CF) and an adolescent short form (TEIQue-ASF). The items on this self-report measure include questions such as "I can tell when a friend is sad" which are answered using a Likert scale. The internal consistency reliability of the test is strong, at 0.84. Further, Mavroveli and Sanchez-Ruiz (2011) found that children who scored higher on trait EI were rated by their peers as significantly more socially competent.

Another measure that has been used is the "Guess Who" peer assessment technique. This technique involves giving students descriptions of certain behaviors that other students may exhibit. Students then list other students' names for whom they think the descriptions are appropriate. Teachers may also provide responses. Scales used in the technique may include categories like "cooperative," "disruptive," "shy," "aggressive," "dependent," "a leader," "intimidating," and "is kind" and "is a bully" (Mavroveli, Petrides, Sangareau, & Furnham, 2009). Students rated as "kind" on this technique showed significantly higher levels of emotional intelligence, whereas students rated as "bullies" showed significantly lower levels of trait EI (Mavroveli et al., 2009)

Empathy There are a variety of interesting methods for measuring empathy found in the empirical literature. One common measure of empathy is the Bryant Index of Empathy Measurement for children and adolescents. This index asks students 22 questions about their emotional state in certain situations, to which they may respond yes or no (Leontopoulou, 2010). The index has also been adapted so that responses are made on a 1–5 Likert scale of "I strongly agree" to "I strongly disagree" (Woods, Wolke, Nowicki, & Hall, 2009). Leontopoulou reported a Cronbach's alpha of 0.52 for the Index, and Woods et al. reported a Cronbach's alpha of 0.70 for the Index.

For younger students, empathy can also be measured with the Southampton Test of Empathy for Preschoolers (STEP). STEP is a test that involves showing students video clips of emotional situations and asking students how both they and the peo-

ple in the video would respond to the situation emotionally. Questions take the format "How does [character's name] feel? How did you feel when [character experienced event]?". Students use images of emotion faces to answer these questions (Howe, Pitten Cate, Brown, & Hadwin, 2008). Howe et al. found a Cronbach's alpha of 0.70 for questions related to understanding how the people in the video would feel (STEP-UND) and 0.86 for sharing an emotional experience with the people in the video (STEP-SHA). In addition, STEP scores were positively correlated with parent evaluations of children's empathy and teacher evaluations of students' prosocial behavior.

Another approach to measuring empathy is based on the self-report methodology. A good example of this approach is the Basic Empathy Scale (BES; Jolliffe & Farrington, 2006). This 20-item scale assesses both cognitive and affective empathy and was designed to measure the degree to which a person understands and shares the emotions of another. Cognitive empathy is tapped by items such as "It is hard for me to understand when my friends are sad," whereas affective empathy is tapped by items such as "I usually feel calm when other people are scared." The BES has been shown to have strong internal consistency reliability (alpha = 0.87 full scale), and confirmatory factor analyses have shown a good fit of the two-factor model. The BES has demonstrated validity by showing positive associations with prosocial behavior and negative correlations with bullying.

Another popular self-report measure related to empathy is the Interpersonal Reactivity Index (IRI) developed by Davis (1983). The IRI includes 28 items that are intended to tap four domains: (1) perspective-taking, (2) fantasy, (3) empathetic concern, and (4) personal distress. Examples of items are "Before criticizing somebody, I try to imagine how I would feel if I were in their place" and "When I see someone being taken advantage of, I feel kind of protective toward them." Participants are to rate each item on a 5-point scale ranging from "Does not describe me well" to "Describes me very well." The instrument has demonstrated strong internal consistency reliability (alpha = 0.80 and higher) as well as strong test–retest reliability (Batanova & Loukas, 2013). Mestre, Frias, and Samper (2004) replicated the four-factor structure of the instrument, thereby providing further evidence in support of its construct validity.

A fourth approach to the measurement of empathy comes from the performance-based tradition. In that regard, the multifaceted empathy test (MET) is exemplary (Dziobek et al., 2008). In this test, 40 photographs depicting different people in positive or negative emotional situations are presented. Cognitive empathy is assessed by asking participants to choose one of four adjectives that best describes the emotional state of the depicted person. For the explicit assessment of affective empathy, participants are asked to rate how strongly they feel for the person on a 9-point rating scale (1 "not at all" and 9 "very strongly"). Participants are also asked to rate on the same 9-point rating scale how strongly they feel affected by the presented photograph in order to assess implicit affective empathy. The MET has strong internal consistency reliability, with the Cronbach's alpha ranging between 0.71 for the cognitive and 0.91 for the explicit emotional empathy scale and 0.92 for the implicit emotional empathy scale. The MET has demonstrated evidence

of construct validity by discriminating between healthy individuals and patients with Asperger's autism (Dziobek et al., 2008). Evidence for the convergent validity of the instrument comes from the fact that scores on the MET show significant correlations with the IRI.

In an approach similar to the MET, Rae Westbury and Neumann (2008) developed a video-based assessment in which they measured empathy physiologically. Participants were shown 10 s video vignettes of five animal groups (humans, primates, companion mammals, utilitarian mammals, and chickens) in distressing situations. Physiological recording sensors for corrugator EMG, skin conductance, and respiration were attached to each participant. After each film clip, participants were prompted on the screen with the written instruction "Please make rating now" to rate their level of empathic feeling using a scale of 0–9 (where 0 = none and 9 = maximal response) using a computer keyboard. Participants were instructed to separate any feelings of disgust from their ratings. Their approach found strong convergent validity with the BEES (Mehrabian, 1996), another psychometrically validated measure of the affective element of empathy.

Finally, as a behavioral measure of empathy, one can measure daily helping. For example, Rameson, Morelli, and Lieberman (2012) used an assessment called the Daily Experience Survey. They asked participants in their study to complete an end-of-day online survey for 14 consecutive days. Two forms of daily helping were measured: stranger–acquaintance helping (e.g., picking up dropped objects and holding a door open; $\alpha = 0.82$) and friend helping (e.g., lending money and giving a ride; $\alpha = 0.73$).

Self-esteem By far the most frequently used measure of self-esteem is Rosenberg's (1965) Self-Esteem Scale. This scale measures general self-esteem and includes ten items capturing a continuum of self-worth statements. The scale has been used extensively in samples with a variety of ages, nationalities, and socioeconomic levels. The internal consistency reliability is strong across studies (alphas range from 0.72 to 0.87). Examples of items are "I feel that I am a person of worth, at least on an equal plane with others" and "I wish I could have more respect for myself" with the latter item being reverse scored.

The Coopersmith Self-Esteem Inventory is widely used to measure self-esteem and is relevant for K-12 students. It has 50 yes/no questions as items, which relate to global self-esteem and relationships with parents and friends. The inventory includes questions like "I am a lot of fun to be with," "I have a low opinion of myself," "I often wish I were someone else," and "Kids usually follow my ideas." Hills, Francis, and Jennings (2011) found that it could be reduced to 19 items while simultaneously yielding improved psychometric properties. Specifically, the internal consistency reliability of the shortened scale is strong (Cronbach's alpha = 0.86), and confirmatory factor analyses validated the three key sources of self-esteem (personal self-esteem, self-esteem derived from peers, and self-esteem derived from parents).

The Piers-Harris Children's Self-Concept Scale (CSCS) has also been used to measure students' self-esteem. The two CSC scales most related to self-esteem are

the happiness and satisfaction scale and the popularity scale. The happiness and satisfaction scale contains 10 items, and the popularity scale contains 12 items. Rousseau, Drapeau, Lacroix, Bagilishya, and Heusch (2005) report that the Cronbach's alpha ranged from 0.57 to 0.71

Whereas the instruments just mentioned attempt to measure global (i.e., domain-general) self-esteem, there are many contexts in which it is more appropriate to measure self-confidence within a particular domain. For this purpose, one psychometrically strong assessment is the Personal Evaluation Inventory (PEI) (Shrauger & Schohm, 1995). This 54-item instrument has eight domain-specific subscales: "academic performance," "physical appearance," "athletics," "romantic relationships," "social interactions," "speaking before others," "general confidence," and "mood" state. The subscales contain between 5 and 7 items each. All items are presented as Likert scales that scored 1–4 (negative items reversed), with 4 indicating strong agreement with items reflecting self-confidence. Examples of items are "I am pleased with my physical appearance" and "I have no difficulty maintaining a satisfying romantic relationship." Evidence for the competency validity of the PEI scores comes from correlational studies showing that PEI scores are significantly correlated in expected directions with other independent measures of anxiety, hopelessness, depression, and optimism (Shrauger & Schohm, 1995). The Cronbach's alpha for the PEI instrument in the original development of this questionnaire ranged from 0.77 for the academic subscale to 0.91 for the athletics subscale. One-month test–retest reliabilities ranged from 0.73 for the academic subscale to 0.90 for the athletics subscale, indicating good stability.

Motivation As with self-esteem, motivation can be thought of and measured in ways that are domain general or ways that are domain specific (see Chap. 10). One of the most prominent measures of domain-general motivation is the Intrinsic Motivation Inventory (IMI) – a self-report inventory designed to assess the level of intrinsic motivation experienced by an individual engaged in an achievement-oriented task (Ryan, Mims, & Koestner, 1983). Six subscales measure various underlying dimensions or indices of intrinsic motivation: (1) interest/enjoyment, (2) perceived competence, (3) effort/importance, (4) pressure-tension, (5) perceived choice, and (6) value/usefulness. All items are scored on a 7-point Likert-type scale ranging from 1 (*not at all true*) to 7 (*very true*). Example items include "I enjoyed doing this activity very much" and "I tried very hard on this activity." The interesting part about the IMI is that it can be adapted to almost any activity. Because motivation is typically believed to be domain specific (i.e., related to particular activity), however, the IMI provides a framework for assessing specific activities. For example, Amorose and Horn (2001) evaluated the psychometric properties of a sport version of the IMI and found that the internal consistency reliability estimates ranged from 0.70 to 0.80. McAuley, Duncan, and Tammen (1987) showed evidence for the construct validity of a higher-order factor of intrinsic motivation with four second-order factors related to specific elements of sport.

In addition, the PISA study measures student motivation within the context of four domains: (1) perseverance (an index derived from items asking students about their willingness to continue working on difficult problems, even when they encounter problems), (2) openness to problem solving, (3) locus of control (the extent to which they see outcomes as being associated with their own effort), and (4) intrinsic vs. instrumental motivation to learn mathematics (http://www.oecd.org/pisa/keyfindings/PISA2012-Vol3-Chap3.pdf).

There are a variety of assessments used to measure student motivation at different levels of their education and in specific domains. For example, the Preschool Reading Attitude Scale (PRAS) and the Emergent Readers Motivation and Reading Scale (ERMAS) tap preschoolers' motivation for reading. The PRAS contains 34 items which students rate with a 1–3 Likert scale of emotion faces (happy, neutral, and sad) (Sperling, Sherwood, & Hood, 2013).

For middle school students, Brookhart, Walsh, and Zientarski (2006) used a collection of scales to measure motivation in social studies and science. These scales included "perceived task characteristics," "perceived self-efficacy," "mastery goal orientations," "performance goal orientations," "amount of invested mental effort," "active learning strategy use," and "superficial learning strategy use." A 1–5 Likert scale ranging from "Yes!" to "No!" was used for student responses.

An assessment of motivation for high school students is the Student Motivation and Engagement Scale-High School (MES-HS). This scale has 44 items which are rated with a 1–7 Likert scale of "strongly disagree " to "strongly agree" (Plenty & Heubeck, 2011). The assessment is typically used to assess motivation in general and, however, can be adapted to assess motivation in a particular subject area, such as math, by rewording questions to pertain to this subject. A similar questionnaire developed by Yin et al. (2008) focuses on motivation in science.

Because motivation is measured in a domain-specific way, there have been other efforts to assess it using specific instruments in the domains of reading, motivation for studying, and even friendship. Logan, Medford, and Hughes (2011) have developed a 15-item questionnaire called the Motivation for Reading Questionnaire (MRQ) that has yielded strong internal consistency (alpha = 0.75). Another interesting measure of motivation comes from Ojanen, Sijtsema, Hawley, and Little (2010). They state that their measure of intrinsic and extrinsic motivation was partly adapted from the Reasons Survey (Ryan & Connell, 1989) and the Achievement Goal Questionnaire (Elliott & Sheldon, 1997) and partly developed by their research team. As Ojanen et al. note, "All motivation items were assessed with respect to three frames: 'Why do you make new friends', 'Why do you get a friend to do something together with you', and 'Why do you keep a good friend'. The participants provided answers in a Likert-scale, ranging from 1 to 5 (1 = I disagree, 5 = I agree). Two items (with respect to the three frames, i.e., six questions altogether) were used to measure intrinsic motivation (alpha = 0.73) and four items (with respect to the three frames, i.e., 12 questions altogether) were used to measure extrinsic motivation (alpha = 0.92). An example of a frame/item combination from the intrinsic scale reads: [Frame] 'Why do you make new friends?' [item] 'Is it because you

enjoy doing it?' The other intrinsic item was: 'Is it because you like to do it?' The extrinsic items were: 'Is it because you want to make your parents happy? Is it because you want to get praise from your teachers? Is it because you don't want your teachers to think you are no good?; Is it because you don't want your parents to be angry with you?'"

Self-directed Learning Historically speaking, the most widely used measure of self-directed learning is Guglielmino's Self-Directed Learning Readiness Scale (1978). The original scale consisted of 41 items that exhibited high internal consistency reliability overall (alpha = 0.81); however, the items were thought to constitute eight dimensions of self-directed learning. Guglielmino reported that principal component analysis (PCA) with varimax rotation yielded an eight-factor structure. She labeled these factors (1) openness to learning opportunities, (2) self-concept as an effective learner, (3) initiative and independence in learning, (4) informed acceptance or responsibility for one's own learning, (5) love of learning, (6) creativity, (7) future orientation, and (8) ability to use basic study skills and problem-solving skills. Response options are on a 5-point Likert scale ranging from "Almost never true of me; I hardly ever feel this way" to "Almost always true of me; there are very few times when I don't feel this way." Example items include "I know what I want to learn," "I don't work very well on my own," and "No one but me is truly responsible for what I learn."

Although the SDLRS is widely used in practice, there is controversy surrounding its psychometric properties. Specifically, advanced data analyses call into question the eight-factor model and appear to demonstrate six highly correlated factors which could be subsumed under a single higher-order factor (West & Bentley, 1990). Other authors have pointed out that the psychometric properties are inconsistent across race and class populations (Straka, 1995). As a result, other researchers (e.g., Abd-El-Fattah, 2010) have created new scales of self-directed learning aptitude. The full scale consists of 40 items and has confirmatory factor evidence for three subscales.

Another sound instrument for measuring self-directed learning is the Self-Directed Learning Scale, which consists of ten items with responses made on a 5-point Likert scale ranging from strongly disagree to strongly agree. Examples of items include "I am good at finding the right resources to help me do well in school" and "If there is something I need to learn, I find a way to do so right away." The SDLS has been found to be an internally consistent measure with Cronbach's alpha values in the mid to high 0.80s (Lounsbury, Saudargas, Gibson, & Leong, 2005). Evidence for the validity of the instrument comes from its positive association to college student life satisfaction and negative association with intention to withdraw from college.

Finally, another scale used to measure students' self-directed learning is the Self-Directed Learning with Technology Scale (SDLTS) developed by (Timothy et al., 2010). This scale includes seven items such as "I go online to ask my teachers questions on my lessons when I am not in school" and "I use the computer to work

with information for my learning", and is intended for use with elementary school students. The results of the SDLTS scale were tested with 398 middle school students, 568 high school students, and 1159 college students. SDLTS scores were significantly predictive of cumulative GPA along with college and life satisfaction. It also exhibited convergent validity with Guglielmino's Self-Directed Learning Readiness Scale.

Summary The competencies chosen to represent the domain of emotional development are certainly not exhaustive, nor are the instruments discussed in this section. However, we have tried to highlight three important features. First, there exist many psychometrically sound instruments for measuring each of these competencies. Second, many of these instruments are specifically tailored to K-12 populations. And third, there are a variety of interesting methods used for assessment. The particular assessments mentioned in this section include some self-report measures, but also many other innovative approaches to measuring emotional development. Table 3.1 presents a summary that summarizes the measures discussed in this section, highlighting the competency they are intended to measure, their psychometric properties, example items, and how to acquire these instruments.

3.5 Assessing Effective Citizenship

Of the 421 randomly sampled public high school mission statements analyzed by Stemler et al. (2011), a total of 58 % of the mission statements endorsed civic development, making it the most highly ranked theme out of 11 major themes identified in their study. In the context of higher education, citizenship was emphasized in 68 % of the statements of essential capabilities of national US universities, making it the third most highly cited essential capability out of ten total capabilities (Stemler, 2012a). Thus, we next review competencies associated with the competency of citizenship. These competencies include (1) civic knowledge, (2) leadership, (3) teamwork, and (4) ethics. At the end of this section, we present a table that summarizes the competency under investigation, the name of the instruments used to assess the competency, example items, their psychometric properties, and information on how to obtain the instrument.

Civic Knowledge Without question, the most widely used and high-stakes test of citizenship is the United States Citizenship and Immigration Services (USCIS) naturalization test. Indeed, this test is a useful indication of what our government considers important in its citizens. In 2007, the assessment was revised and shifted its emphasis more toward larger concepts of American democracy and the rights and responsibilities of American citizens rather than focusing on particular historical facts. For example, a question that is formerly asked, "What country did we fight during the Revolutionary War?", has been revised to read "Why did the colonists fight the British?". Because this test represents the American government's definition of effective citizenship, its structure and content are important to

Table 3.1 Selected measures of emotional intelligence, empathy, self-esteem, motivation, and self-directed learning

Measure	Sample item/item types	Psychometrics	Availability
Emotional intelligence			
Bar-On EQ-i youth version	"When I need to get a job done, I can get myself into the right emotional state to get it done"	Solid levels of internal consistency reliability ranging from 0.69 to 0.86 and test–retest reliability coefficients ranging from 0.55 to 0.82. Shown to predict academic success	Purchase through MHS (http://www.mhs.com)
MSCEIT	Some items have participants view a face and rate the extent to which different emotions are present or absent from the facial expression. Other items are scenario based and ask participants what emotions would be most useful in the given scenarios	High levels of internal consistency reliability (alpha=0.91 full scale) as well as strong test–retest reliability (alpha=0.86 full scale) Has been shown to predict a variety of important outcomes, including job performance and academic success	Purchase through MHS (http://www.mhs.com)
STEM-Y	"You and James sometimes help each other with homework. After you help James on a difficult project, the teacher is very critical of this work. James blames you for his bad grade. You respond that James should be grateful, because you were doing him a favor. What would you do in this situation?"	Significant predictor of GPA and problem-focused coping	Contact: C. MacCann carolyn.maccann@sydney.edu.au
TEIQue-CF	"I can tell when a friend is sad"	Internal consistency of 0.84 Convergent validity evidence with peer ratings of social competence	Contact: K. V. Petrides k.petrides@ucl.ac.uk
"Guess Who" peer assessment	Students nominate peers who fit certain descriptions (e.g., cooperative, kind, a bully)	Convergent validity evidence between trait EI and ratings of kindness (positive) and bullying (negative)	Available in Mavroveli et al. (2009)
Empathy			
Bryant Index of Empathy Measurement for children and adolescents	Asks questions about their emotional state in certain situations to which they respond Yes or No. The scale has also been modified to a Likert scale (strongly agree to strongly disagree)	Reported Cronbach's alphas range from 0.52 to 0.70 for the Index	de Wied et al. (2012)

Southampton Test of Empathy for Preschoolers (STEP)	"How does [character's name] feel?" "How did you feel when [character experienced event]?"	Cronbach's alpha of 0.70 for questions related to understanding how the people in the video would feel and 0.86 for sharing an emotional experience with the people in the video STEP scores were positively correlated with parent evaluations of children's empathy and teacher evaluations of students' prosocial behavior	Howe et al. (2008)
Basic Empathy Scale	"It is hard for me to understand when my friends are sad" "I usually feel calm when other people are scared"	Cronbach's alpha of 0.86 (full scale); confirmatory factor analyses have shown a good fit of the two-factor model Results have shown convergent validity with prosocial behavior and negative correlations with bullying	Free of charge for research purposes by writing to D. Jolliffe at dj211@cam.ac.uk
Interpersonal Reactivity Index (IRI)	"Before criticizing somebody, I try to imagine how I would feel if I were in their place" "When I see someone being taken advantage of, I feel kind of protective toward them"	Strong internal consistency reliability (alpha = 0.80 and higher) as well as strong test–retest reliability	Available from the original publications (Davis, 1980, 1983)
Multifaceted Empathy Test (MET)	Participants are shown photographs depicting people in positive or negative emotional situations. Participants are asked to rate how strongly they feel for the person in the picture	Cronbach's alpha ranging between 0.71 for the cognitive and 0.91 for the explicit emotional empathy scale and 0.92 for the implicit emotional empathy scale Scores on the MET show convergent validity with the IRI	Available from the first author, I. Dziobek, at: dziobek@mpib-berlin.mpg.de
Daily Experience Survey	Stranger–acquaintance helping (e.g., picking up dropped objects and holding a door open) and friend helping (e.g., lending money and giving a ride)	Stranger–acquaintance helping Cronbach's alpha 0.82; friend helping Cronbach's alpha 0.73	Rameson, Morelli, and Lieberman (2012)
Self-esteem			
Rosenberg's (1965) Self-Esteem Scale	"I feel that I am a person of worth, at least on an equal plane with others" "I wish I could have more respect for myself"	Cronbach's alpha ranges from 0.72 to 0.87 across studies	Available from the original publication (Rosenberg, 1965) and easily found online

(continued)

Table 3.1 (continued)

Measure	Sample item/item types	Psychometrics	Availability
Coopersmith Self-Esteem Inventory	"I am a lot of fun to be with" "I have a low opinion of myself" "I often wish I were someone else" "Kids usually follow my ideas"	Cronbach's alphas of 0.86 for the reduced, 19-item scale Construct validity evidence for three-factor solution (personal self-esteem, self-esteem derived from peers, self-esteem derived from parents)	Available in Hills, Francis, and Jennings (2011) article
Piers-Harris Children's Self-Concept Scale (CSCS)	The happiness and satisfaction scale and the popularity scale are most closely aligned with self-esteem	Cronbach's alpha ranges from 0.57 to 0.71 across studies	Contact: K. V. Petrides k.petrides@ucl.ac.uk
Personal Evaluation Inventory (PEI)	"I am pleased with my physical appearance" "I have no difficulty maintaining a satisfying romantic relationship"	Cronbach's alpha in the original development of this questionnaire ranged from 0.77 for the academic subscale to 0.91 for the athletics subscale. One-month test–retest reliabilities ranged from 0.73 for the academic subscale to 0.90 for the athletics subscale PEI scores are significantly correlated in expected directions with other independent measures of anxiety, hopelessness, depression, and optimism	Available in Robinson, Shaver, and Wrightsman (1991) and by searching online
Motivation			
Intrinsic Motivation Inventory (IMI)	"I enjoyed doing this activity very much" "I tried very hard on this activity"	Internal consistency reliability estimates ranged from 0.70 to 0.80	Freely available for research purposes from: www.ravansanji.ir/?std1019IMIfull
PISA Drive and Motivation scale	"I put off difficult problems" "I like to solve complex problems" "If I put in enough effort, I could succeed in mathematics"	No validity or reliability information is reported; however, the items are derived from well-known scales	Available from PISA report at: www.oecd.org/pisa/keyfindings/PISA2012-Vol3-Chap3.pdf

Motivation for Reading Questionnaire (MRQ)	"If a book is interesting, I don't care how difficult it is to read" "I like reading so that I can learn more about things" "Time goes very slowly when I am reading"	Cronbach's alpha 0.75	Available from the Appendix of Logan, Medford, and Hughes (2011)
Ojanen, Sijtsema, Hawley, and Little's (2010) measure of intrinsic and extrinsic motivation	"Why do you make new friends?" Is it because you enjoy doing it? Is it because you like to do it? Is it because you want to make your parents happy? Is it because you want to get praise from your teachers? Is it because you don't want your teachers to think you are no good? Is it because you don't want your parents to be angry with you?	Cronbach's alphas range from 0.73 to 0.92 by scale (intrinsic vs. instrumental)	Available from Tiina Ojanen, at: tojanen@usf.edu
Self-directed learning			
Guglielmino's Self-Directed Learning Readiness Scale	"I know what I want to learn" "I don't work very well on my own" "No one but me is truly responsible for what I learn"	High internal consistency reliability overall (Cronbach's alpha 0.81). Principal component analysis with varimax rotation yielded an eight-factor structure	Available from Learning Preference Assessment, at: http://www.lpasdlrs.com
Self-Directed Learning Scale	"I am good at finding the right resources to help me do well in school" "If there is something I need to learn, I find a way to do so right away"	Cronbach's alpha values in the mid- to high 0.80s Positive association to college student life satisfaction and negative association with intention to withdraw from college	Available in Lounsbury, Levy, Park, Gibson, and Smith (2009)
Self-Directed Learning with Technology Scale (SDLTS)	"I go online to ask my teachers questions on my lessons when I am not in school" "I use the computer to work with information for my learning"		Available in Timothy et al. (2010) article

keep in mind when thinking about what civic skills should be expected of students and how best to measure them. Curiously, no psychometric information on the test is publicly available; however, the test is used for granting naturalization to citizens. Importantly, the state of Arizona has just passed legislation requiring high school students to pass the US naturalization test in order to receive a high school diploma (Rojas & Rich, 2015). Several other states currently have similar legislation pending.

From a scientific perspective, the gold standard for assessing citizenship comes from the National Assessment of Educational Progress (NAEP). NAEP tests students in grades 4, 8, and 10 in three areas: (1) civic knowledge, (2) intellectual and participatory skills, and (3) civic dispositions. The following are some example items: "Identify a civic responsibility (e.g., voting)"; "What is the main source of government funding (taxes)?"; "Why does the US Constitution limit the powers of government?"; "Identify one way to express an opinion on a public policy issue"; "Identify the appropriate official to contact to solve a problem." Not all items are released to the public; however, subsets of items can be found within their technical manual (Carr, 2014). The test is scored using item response theory and scale anchoring (one assumes from the scant reporting in the technical manuals); however, no psychometric data regarding fit indices for items are available publicly, nor are any validation data. Remarkably, however, the NAEP civic assessment has been suspended indefinitely for fourth and twelfth graders due to sequestration (Klein, 2013) – only eighth graders will continue to take the test. Notably, no cuts were made to cognitively oriented assessments.

The International Association for the Evaluation of Educational Achievement (IEA) Civic Education Study (International Association for the Evaluation of Educational Achievement, 1999) was an international assessment of the civic knowledge and skills of youths around the world, and its goal was to investigate, using a comparative framework, the ways in which youths are prepared for their roles as citizens in democracies and aspiring democracies. In the late 1990s, the test was administered to a group of about 90,000 14-year-olds from 28 countries and a group of about 60,000 16–18-year-olds from 16 countries. The test contained three sections: (i) a multiple-choice section that tested civic knowledge and skills in interpreting civic information, (ii) a section asking for background information, and (iii) a section on conceptual understanding, attitudes, behavior, and actions. This last section was further divided to include civic knowledge, interpretation skills, economic literacy, conventional citizenship, social movement-related citizenship, trust in governmental institutions, economy-related government responsibilities, society-related government responsibilities, positive attitude toward one's nation, positive attitude toward immigrants, confidence in participation in school, expected participation in political activities, and open climate for classroom discussion. Cronbach's alphas ranged from 0.57 to 0.86 for the various subscales, but no validity information has been published.

Stemler (2015) has recently developed a new test of citizenship that is targeted at K-12 students. The test is designed to measure civic attitudes, behaviors, and

cognitions (knowledge). The test was administered to 442 K-12 students, and the test demonstrated strong construct validity as well as predictive validity with teachers' ratings of the students' levels of civic engagement.

Leadership One of the most widely used measures of leadership is the Multifaceted Leadership Questionnaire (MLQ) developed by Bass and Avolio (1994). The instrument can be used as a self-report questionnaire, but can also be used in the form of a 360° rating scale assessment by asking others to evaluate a person of interest with regard to each of the items on the rating scale. In this way, multiple perspectives on the individual's leadership style emerge. Bass and Avolio have conceptualized leadership as falling into one of three main forms: (1) *transformational* (in which leaders inspire others, motivate action, and challenge those around them intellectually), (2) *transactional* (in which leaders manage those around them by contingent rewards), and (3) *passive/avoidant* (in which leaders manage by exception, i.e., pointing out mistakes of the subordinates, and generally take a passive role. The MLQ rating scale consists of 45 questions that are to be responded to on a Likert scale where 0 = "Not at all," 1 = "Once in a while," 2 = "Sometimes," 3 = "Fairly often," and 4 = "Frequently, if not always." Examples of questions include "Talks optimistically about the future," "Avoids making decisions," and "Spends time teaching and coaching." Various forms of the MLQ have been developed, including a shorter form. The internal consistency reliability of the instrument is excellent, with Tejeda, Scandura, and Pillai (2001) reporting values of the subscales ranging from 0.69 to 0.92, with a median value of 0.88 across four different samples. In addition, the instrument has been shown, from studies reported in the user manual as well as independent investigations, to have strong competency validity.

The Roets Rating Scale for Leadership (RRSL) has also been used to measure student leadership. It includes 26 items, which students use to provide a self-report rating of their leadership. Examples of items include "Listen to both sides," "Think one can do well as a leader," "Can work with different person types," and "Can say opinions in public" (Chan, 2000). Similarly, the leadership subscale of the Scales for Rating the Behavioral Characteristics of Superior Students (SRBCSS) has also been used to measure student leadership. This subscale includes seven items, and parent and teacher ratings on these items are used to evaluate students' leadership. He found that RRSL and SRBCSS leadership scores were significantly correlated with each other. Teacher ratings of student leadership on the Behavioral Assessment System for Children (BASC) have also been used as a measurement of student leadership with a reliability of 0.83 (Tolan & Larsen, 2014).

There also exist measures of leadership for teachers and principals. One of these measures is the Vanderbilt Assessment for Leadership in Education (VAL-ED). The VAL-ED is considered a "360° assessment" (Covay Minor et al., 2014) because it has teachers and supervisors rate their school principal in 36 domains of leadership. Ratings are done on a 1 to 5 Likert scale, ranging from "ineffective" to "outstandingly effective" (Covay Minor et al., 2014). They reported that the VAL-ED produces ratings of principals' effectiveness that exhibit 70 % agreement with ratings of the same principals' performance provided by school superintendents. Other

measurements that can be used are measures of principal instructional leadership and teacher–principal trust (Sebastian & Allensworth, 2012). These measures involve ratings from teachers and principals and have been shown to have a reliability of 0.91 and 0.89, respectively.

Teamwork and Cooperation As with many of the instruments we have reviewed, we begin here by summarizing an instrument that historically has been used quite frequently to measure teamwork. The Teamwork Knowledge Skill and Ability (TKSA) Test is one measure that has been widely used to assess participants' interpersonal and self-management knowledge (Stevens & Campion, 1999). The instrument includes 35 multiple-choice items. An example item is as follows: "Suppose that you find yourself in an argument with several coworkers about who should do a very disagreeable but routine task. Which of the following would likely be the most effective way to resolve this situation? The four response options for this question are: (A) Have your supervisor decide, because this would avoid any personal bias. (B) Arrange for a rotating schedule so everyone shares the chore. (C) Let the workers who show up earliest choose on a first-come, first-served basis. (D) Randomly assign a person to do the task and don't change it."

Although the test authors (Stevens & Campion, 1999) have reported internal consistency reliability as high as 0.80, several other authors have reported lower consistency reliability estimates (e.g., alpha = 0.37–0.59). Furthermore, the primary authors contend that the instrument measures five dimensions of teamwork; however, confirmatory factor analyses suggest that single overarching factor fits the data best.

Consequently, Aguado, Rico, Sanchez-Manzanares, and Salas (2014) recently developed the Teamwork Competency Test (TCT), which was inspired by TWKSAT but that is designed to improve upon its psychometric properties. The TCT consists of 36 items that describe different situations that may arise within a work team. Respondents are asked to use a 4-point Likert scale, where "0" = never/almost never and "4" = always/almost always. Example items are "I often get involved in monitoring the task performance of other team members," "I care and act to make team conflicts explicit in a way that they can be solved," and "I provide my peers with relevant information on how well I think the team tasks are progressing." The TCT has excellent evidence supporting internal consistency reliability (alpha = 0.84 full scale). Furthermore, the instrument has strong predictive validity evidence associated with supervisor and self-evaluations of team performance.

The Individual Performance in Teams Scale (IPIT) is another popular measure of teamwork. It uses a rating-scale approach to the assessment of performance in teams. The IPIT consists of 33 items depicting various aspects of team member behavior related to conflict resolution, collaborative problem solving, communication, performance management, and task coordination. The items reflect issues/topics that previous literature has found to be important for team functioning. Sample behavior items include statements such as "tried to keep group aware of time issues," "responded calmly to others," and "helped resolve any conflicts." The IPIT was designed to be used by raters viewing videotaped team interactions. Raters are

asked to indicate on a 7-point scale of behavior frequency with anchors ranging from 1 (to no extent) to 7 (to a great extent) the extent to which the behavior was observed. A "not applicable" was also an option. Items thought to impede team performance (e.g., rudely interrupted other members) were reverse coded. The scale also contains one general item concerning the team member's effectiveness during the team task. While interrater agreement should be assessed with each new study, McClough and Rogelberg (2003) found in their study that the three raters they used achieved strong levels of interrater reliability (intraclass correlation = 0.73). Peer ratings of individual teamwork performance were also gathered. Team members were asked to rate each other on five dimensions (e.g., participation in the group, interpersonal skills). However, the scale did not correlate, in that study, with peer ratings of team performance. This is a common danger associated with 360° feedback approaches – self- and other reports are rarely in alignment (Atwater, Ostroff, Yammarino, & Fleenor, 1998).

Similarly, in the classroom, situational judgment tests (SJTs) have been used to assess teamwork, in addition to self-report and teacher rating scales (Wang, MacCann, Zhuang, Liu, & Roberts, 2009). The SJT assessment gives students a scenario involving a group situation and asks them to rate how effective four possible responses to the given scenario would be. Students' effectiveness ratings are then used to calculate their SJT score. SJT scores were shown to correlate with teacher ratings of teamwork as well as self-report measures. In the Wang et al. study, all three methods showed convergent validity, but only teacher ratings of teamwork predicted composite course grades.

Ethics Perhaps the most widely used domain-general test of moral reasoning is the *Defining Issues Test* (Rest, 1979; Rest, Narvaez, Bebeau, & Thoma, 1999), which is based on Kohlberg's (1969) theory of moral development. Building on Piaget's work on moral judgment, Kohlberg advanced a theory of moral development comprising six stages, divided into three levels: (i) pre-conventional morality, (ii) conventional morality, and (iii) post-conventional morality. The *Defining Issues Test* consists of five situational dilemmas to which respondents are asked to respond. The classic example is the Heinz dilemma in which Heinz's wife is dying of cancer and a chemist has discovered a cure but is charging ten times more than Heinz can afford to pay and he will not alter the price. Without the drug, the wife will die. Participants are asked not only to select what they think Heinz should do (i.e., steal the drug or not) but also to rate the importance of various reasons for their decision (e.g., because it is against the law to steal; because he loves his wife and she will die without the medicine, so it is worth breaking the law). Kohlberg's theory has been criticized for its lack of applicability across gender (Gilligan, 1982) and cultures (Tietjen & Walker, 1985) and for the fact that it does not strongly correlate with actual behavior (Blasi, 1980; Gibbs et al., 1986). Item total correlations range from 0.48 to 0.77 with the internal consistency reliability for the overall scale reported to be 0.65 (Xu, Iran-Nejad, & Thoma, 2007).

An alternative conception of ethical thinking that overcomes some of the criticisms of the Kohlberg model has been advanced by Schwartz (1992). According to

Schwartz, there are ten universal ethical values (power, achievement, hedonism, stimulation, self-direction, universalism, benevolence, tradition, conformity, and security) that have been demonstrated to hold up empirically across more than 50 cultures (Schwartz & Bardi, 2001) and that relate, at least modestly, to relevant behavioral outcomes (Schwartz, 2007). The *Schwartz Value Scale (SVS)* is a 56-item self-report instrument that measures the extent to which individuals endorse different ethical values as guiding principles in their lives. Participants are asked to rate, on a scale ranging from 0 (opposed to my principles), 1 (not important), 4 (important), to 8 (of supreme importance), the importance as a life-guiding principle for them terms such as "power, that is, social power, authority, and wealth," and "achievement, that is, success, capability, ambition, and influence on people and events." Internal consistency reliabilities for subscales have been reported on the order of 0.58 and 0.60.

Recently, Stemler (2012b) has proposed an alternative instrument called the *Ethical Priority Test* that represents a compromised position between the universalism advocated by Kohlberg/Rest and the relativism advocated for by Schwartz. Specifically, like Schwartz, Stemler proposes a finite set of universal human values (e.g., honesty, kindness, responsibility, justice). These are basic values that are articulated in almost all religions and are found pervasively in the ethics and philosophy literature. Most people use these values as guiding principles in their lives. When given a choice between honesty and non-honesty, for example, most rational people will choose to be honest. However, Stemler argues that the reason we observe individual differences in behaviors is because reasonable people may prioritize each of the values differently. The test is designed for use with adolescents and college-aged students. An example item follows: "You have recently been appointed captain of your dance team. Earlier today, you fell down the stairs and broke your ankle. When you went to see your doctor, she said that you should rest as much as you can and not participate in the competition tomorrow. If you do not compete, not only will your team be losing one of its best dancers, but they also will need to change all of their formations. Do you decide to compete? (A) Yes; (B) No." A response of "Yes" corresponds to the theoretical value of "responsibility," whereas a response of "No" corresponds to a theoretical value of "safety."

The novelty of this assessment is that it does not yield a single score, but rather a profile of ethical values. Individual profiles (patterns of values) are generated using Latent Class Analysis. The instrument has shown sufficient evidence of test–retest reliability (ranging from 0.47 to 0.69 across scales). In addition, the EPT has demonstrated excellent convergent validity with predicted elements of the Schwartz and Kohlberg scales as well as discriminant validity with personality traits.

A few measures of ethics have been used specifically with students in school settings. One of these measures is the Ethical Sensitivity Scale Questionnaire. The questionnaire has 28 items that encompass the following seven categories: "(1) reading and expressing emotions, (2) taking the perspectives of others, (3) caring by connecting to others, (4) working with interpersonal and group differences, (5) preventing social bias, (6) generating interpretations and options, and (7) identifying the consequences of actions and options." Within these categories, each item is

rated on a 1–5 Likert scale, "totally disagree" to "totally agree" (Tirri & Nokelainen, 2007). Examples of items include "I think it is good that my closest friends think in different ways," "I believe there are several right solutions to ethical problems," "I notice if someone working with me is offended by me," and "I try to consider other people's needs even in situations concerning my own benefits" (Tirri & Nokelainen, 2007). Internal consistency reliability estimates ranged from 0.50 to 0.78 for the seven scales, which is moderate.

An interesting self-report scale of ethical behavior in leisure is the Aristotelian Ethical Behavior in Leisure Scale (AEBLS) (Widmer, Ellis, & Trunnell, 1996). The AEBLS is a 62-item summative scale derived from research with adolescents. Each item represents one of four domains of ethical behavior that are derived in an interpretation of Aristotle's Nicomachean ethics philosophy: intellectual activity, creative activity, moral behavior, and meaningful relationships. Widmer, Ellis, and Munson (2003) created a short form of the scale in which 26 items were selected for inclusion (AEBLS-S). The response format is scaled from one to five: one = never, two = seldom, three = sometimes, four = often, and five = always. Six items are reverse coded. Examples of items include "I think about world problems in my free time," "I spend my holidays with people who are important to me," and "I am fair when I play games." The instrument has strong internal consistency reliability (0.98) and some evidence to support its validity.

Summary As was the case with emotional development, the competencies chosen to represent the citizenship are certainly not exhaustive, nor are the instruments discussed in this section that are used to measure these competencies. However, we have again tried to highlight instruments that are (1) psychometrically sound and (2) specifically tailored to K-12 populations and (3) represent a variety of interesting methods used for assessment. Table 3.2 presents a summary that summarizes the measures discussed in this section, highlighting the competency they are intended to measure, their psychometric properties, example items, and how to acquire these instruments.

3.6 Policy Implications of Aligning Mission and Measurement

Our fundamental thesis in this chapter is that mission and measurement should be aligned. Thus, our first task was to examine the mission of schools. A review of the empirical research drawn from a wide variety of data sources and a broad range of constituents reveals that emotional development and civic development are at least on equal footing with the cognitive element as core purposes of schooling (Stemler et al., 2011; Stemler & Bebell, 2012; Stemler, 2012a). Indeed, the data clearly show that broader competencies are not "add-ons" that are secondary to the cognitive purpose of schooling, but rather that these broader competencies tend to be mentioned as the core purposes of schooling with even greater frequency than the cognitive elements.

Table 3.2 Selected measures of civic knowledge, leadership, teamwork, and ethics

Measure	Sample item/item types	Psychometrics	Availability
Civic knowledge			
United States Citizenship and Immigration Services (USCIS) naturalization test	"Why did the colonists fight the British?"	Not found	Freely available from the United States Citizenship and Immigration Services website (http://www.uscis.gov)
National Assessment of Educational Progress (NAEP)	"Identify a civic responsibility (voting)" "What is the main source of government funding (taxes)?" "Why does the US Constitution limit the powers of government?" "Identify the appropriate official to contact to solve a problem"	Not found	Freely available from: http://nces.ed.gov
IEA Civic Education Test	Civic knowledge Civic behaviors	Cronbach's alpha ranges from 0.57 to 0.86 for various subtests	Sample items available from: http://www.iea.nl/cived.html
Stemler's ABC Citizenship Assessment	"What does being a good citizen of your school look like?" "Draw a picture and describe it" "If you disagreed with an action taken by your class council, what would you do?" "How important to you is it to help make decisions for your neighborhood or town?" "Who is the governor of your state?"	Convergent validity with teachers' ratings of students' civic engagement	Available from steven.stemler@wesleyan.edu
Leadership			
Multifaceted Leadership Questionnaire (MLQ)	"Talks optimistically about the future" "Avoids making decisions" "Spends time teaching and coaching"	Cronbach's alpha of the subscales ranges from 0.69 to 0.92, with a median value of 0.88 across four different samples	Available for purchase through: www.mindgarden.com

Roets Rating Scale for Leadership (RRSL)	"Listen to both sides" "Think one can do well as a leader" "Can work with different person types" "Can say opinions in public"	Cronbach's alpha of 0.83 for the BASC Convergent validity with scales for rating the behavioral characteristics of superior students	Available from Roets (1986)
Vanderbilt Assessment for Leadership in Education (VAL-ED)	A 360° feedback rating scale system where principals, teachers, and superintendents are rated	Internal consistency reliability reported at 0.89–0.91 Convergent validity with superintendent's ratings of teacher leadership	Available for purchase from: http://www.valed.com
Teamwork			
Teamwork, Knowledge, Skill, and Ability (TKSA) Test	Suppose that you find yourself in an argument with several coworkers about who should do a very disagreeable but routine task. Which of the following would likely be the most effective way to resolve this situation? (A) have your supervisor decide, because this would avoid any personal bias; (B) arrange for a rotating schedule so everyone shares the chore; (C) let the workers who show up earliest choose on a first-come, first-served basis; (D) randomly assign a person to do the task and don't change it	Internal consistency reliability as high as 0.80; although several others have reported lower consistency reliability estimates (e.g., alpha = 0.37–0.59)	Available by contacting Michael J. Stevens at: http://kozaigroup.com/who-we-are/michael-j-stevens-ph-d/
Teamwork Competency Test (TCT)	"I often get involved in monitoring the task performance of other team members" "I care and act to make team conflicts explicit in a way that they can be solved" "I provide my peers with relevant information on how well I think the team tasks are progressing"	Cronbach's alpha of 0.84 (full scale)	Available in the Appendix of the Aguado, Rico, Sánchez-Manzanares, and Salas (2014) paper

Table 3.2 (continued)

Measure	Sample item/item types	Psychometrics	Availability
Individual Performance in Teams Scale (IPIT)	"tried to keep group aware of time issues" "responded calmly to others" "helped resolve any conflicts"	Interrater reliability was acceptable (intraclass correlation = 0.73)	Available in the paper by McClough and Rogelberg (2003)
The multimethod assessment of teamwork	Includes self-report teamwork items (e.g., I enjoy sharing ideas) Situational judgment test items Behaviorally anchored teacher rating scale	Factor analyses confirmed the three-factor structure of the self-report measure with Cronbach's alphas ranging from 0.78 to 0.88. Alpha for the SJT was 0.71, while alpha for the teacher scale was 0.98	Contact Richard D. Roberts: rroberts@proexam.org
Ethics			
Defining Issues Test (DIT)	Heinz's wife is dying of cancer and a chemist has discovered a cure but is charging ten times more than Heinz can afford to pay and he will not alter the price. Without the drug, the wife will die. Participants are asked not only to select what they think Heinz should do but also to rate the importance of various reasons for their decision	Item total correlations range from 0.48 to 0.77 with an internal consistency reliability for the overall scale of 0.65	Available from the University of Alabama's Center for the Study of Ethical Development (http://ethicaldevelopment.ua.edu). Costs: approximately $2 per individual to administer and score

Schwartz Value Scale (SVS)	Power, that is, social power, authority, and wealth Achievement, that is, success, capability, ambition, and influence on people and events	Internal consistency reliabilities for subscales on the order of 0.58 and 0.60	Contact: crossculturalcentre@yahoo.com
Ethical Priority Test	Your little sister spent a great deal of time fixing her hair and makeup before her first school dance, and she asks you how she looks. You think she looks ridiculous but you realize that her date will be arriving any minute and telling her what you think might undermine her self-confidence. Do you tell her? (A) Yes; (B) No	Test–retest reliability ranging from 0.47 to 0.69 across scales. Convergent validity evidence with the DIT and SVS and discriminant validity evidence with personality measures	Freely available from: steven.stemler@wesleyan.edu
Ethical Sensitivity Scale Questionnaire	"I think it is good that my closest friends think in different ways" "I believe there are several right solutions to ethical problems" "I notice if someone working with me is offended by me"	Internal consistency estimates range from 0.50 to 0.78 for the seven scales	Contact Kirsi Tirri: Kirsi.Tirri@helsinki.fi
Aristotelian Ethical Behavior in Leisure Scale (AEBLS)	"I think about world problems in my free time" "I spend my holidays with people who are important to me" "I am fair when I play games"	Internal consistency reliability of 0.98	Freely available in the publication (Widmer et al., 1996)

This then led us to question why accountability systems tend to rely almost exclusively on measures of the cognitive domain. We believe that the reason for this is because there is a common misperception that there exist no objective, quantitative ways to measure broader, noncognitive competencies. Thus, the bulk of this chapter was devoted to reviewing psychometrically sound instruments measuring a wide variety of noncognitive competencies that are explicitly valued by businesses, courts, citizens, and schools. Our review was not intended to be exhaustive and our listings could easily be expanded. However, our main objective was simply to illustrate that instruments meeting strong psychometric standards do exist for the broader competencies schools care about. Furthermore, there are many innovative methods for assessment that go well beyond simple self-report measures. Further, a few larger efforts have recently emerged to help fill the gap of providing novel measures of key elements of school mission. Roberts and colleagues have developed what they call the Mission Skills Assessment (http://indexgroups.org/msa/). Stemler and Bebell have developed a number of new measures of broader outcomes aligned with school mission and have collected information on other measures that can be used to measure school mission (see http://www.purposeofschool.com). Consequently, we see no reason why measures of broader, noncognitive skills should be entirely ignored by accountability systems.

Accountability needs to start with the vision of the school – the mission set by the leadership in conjunction with input from the community and especially the teachers working there. Schools should be given the power to determine for themselves what they believe are their most important aims. They should then be given the tools to observe for themselves whether they are making progress toward the aims they hold most dear. When people are allowed to pursue with vigor those things about which they are truly passionate, and when they are supported in their quest, then we will truly see schools that are preparing students to be successful.

We propose that one approach for realizing this vision would be to have instruments, such as those listed in this chapter, included in an "approved list" by the US Department of Education. Instruments that are candidates for inclusion could be submitted to a special committee of the Department of Education for review. This committee would consist of experts in education and psychometrics, and their responsibility would be to determine whether the submitted instruments meet the federal seal of approval for use in measuring progress toward the measurement of a particular skill or outcome. Thus, the federal government would have a list of "approved" measures for a wide variety of important competencies that schools may wish to measure. Schools would then be responsible for articulating the aims they wish to pursue, and the state and federal government would assess progress toward achieving those aims by using one of the "approved" measures. This approach would allow schools to have the autonomy to set their own objectives that are responsive to local needs, while at the same time preserving the external desires for accountability.

The focus would not be exclusively on the extent to which schools had achieved their aims but also on the quality of their instructional plan for making progress. Schools would be required to select aims they care about, collect data on them, and evaluate how they are working and what they might want to do differently or whether they feel a change is needed. These reports would be evaluated every 3–4 years for coherence and accountability purposes. This model of accountability is similar to an accreditation approach and is much more appropriate for school level accountability than simply aggregating individual student scores on tests of cognitive achievement. In the end, our argument is simple. Give schools the choice to determine their goals, which they already do via their mission statement. Let schools develop a plan for what they will do and how they will know that they are accomplishing it. Assist them in this effort by providing a federally reviewed and approved list of instruments they can use to measure whatever competencies they care about. Then, give them time to execute their plan and take a closer look to see whether they have accomplished their aims. If they are making progress, then accolades would follow. If they are not, then further support and/or accountability recommendations would be triggered.

3.7 Conclusion

Aligning measurement with mission is an intuitive concept that is easily grasped by educators, students, parents, businesses, and policy makers. Indeed, it is remarkable that accountability systems could be based on any other premise! We believe that the misalignment of accountability systems and school purpose has been the chief source of resistance to the accountability movement. Most teachers we have met are in favor of the concept of accountability; but they want systems that are aligned with the goals they value. As we have demonstrated previously, a wide variety of sources converge on the importance of emotional development, civic development, and cognitive development as key outcomes of schooling. And there is abundant evidence to suggest that these three competencies are of equal value – in no source of evidence do we find any suggestion that one of the three competencies is of greater worth than the other two. We are not suggesting that an emphasis on the cognitive domain be abandoned. Rather, our goal in this chapter was to point out that (1) there is widespread consensus across a broad range of constituencies about the goals/competencies that form the core purposes of schooling and (2) there is a solid tradition of objective, quantitative, and psychometrically sound assessment of a wide variety of noncognitive competencies. The alignment of mission and measurement need not be a fantasy. It is a realistic option that is well within our grasp. It can and should be the basis for a new sort of accountability system – one in which schools have agency and the federal and state role is to support and monitor progress toward those goals.

References

Abd-El-Fattah, S. M. (2010). Garrison's model of self-directed learning: Preliminary validation and relationship to achievement. *The Spanish Journal of Psychology, 13*(2), 586–596.

AdvanceEd. (2006). *Accreditation standards for quality schools* (No. LC# 2006905052). Decateur: AdvanceEd.

Aguado, D., Rico, R., Sánchez-Manzanares, M., & Salas, E. (2014). Teamwork competency test (TWCT): A step forward on measuring teamwork competencies. *Group Dynamics: Theory, Research, and Practice, 18*(2), 101–121.

Amorose, A. J., & Horn, T. S. (2001). Pre-to post-season changes in the intrinsic motivation of first year college athletes: Relationships with coaching behavior and scholarship status. *Journal of Applied Sport Psychology, 13*(4), 355–373.

Association of American Colleges and Universities. (2010). *Raising the bar: Employer's views on college learning in the wake of the economic downturn*. Washington, DC: Hart Research Associates.

Atwater, L. E., Ostroff, C., Yammarino, F. J., & Fleenor, J. W. (1998). Self-other agreement: Does it really matter? *Personnel Psychology, 51*, 577–598.

Bar-on, R. (1997). *EQ-I technical manual*. Toronto, ON: Multi-Health Systems.

Bass, B. M., & Avolio, B. J. (Eds.). (1994). *Improving organizational effectiveness through transformational leadership*. Thousand Oaks, CA: Sage.

Batanova, M., & Loukas, A. (2014). Unique and interactive effects of empathy, family, and school factors on early adolescents' aggression. *Journal of Youth and Adolescence, 43*(11), 1890–1902.

Bebell, D., & Stemler, S. (2002, April). *MCAS and the purpose of school*. Paper presented at the annual meeting of the New England Educational Research Organization, Northampton, MA.

Bebell, D., & Stemler, S. E. (2004, April). *Reassessing the objectives of educational accountability in Massachusetts: The mismatch between Massachusetts and the MCAS.* Paper presented at the American Educational Research Association, San Diego, CA.

Berleur, J., & Harvanek, R. F. (1997). *Analysis of mission statements or similar documents of Jesuit universities and higher education institutions*. Retrieved October 17, 2002, from http://www.info.fundp.ac.be/~jbl/mis-stat/index.htm

Birney, D., & Stemler, S. E. (2007). Intelligence quotient. In N. Salkind (Ed.), *Encyclopedia of measurement and statistics* (Vol. 2, pp. 473–476). Thousand Oaks, CA: Sage.

Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin, 88*, 1–45.

Brookhart, S. M., Walsh, J. M., & Zientarski, W. A. (2006). The dynamics of motivation and effort for classroom assessments in middle school science and social studies. *Applied Measurement in Education, 19*(2), 151–184.

Carr, P. (2014). *Grade 8 sample questions: Civics, geography, US history, science*. National Assessment of Educational Progress. Downloaded May 28, 2015 from: https://nces.ed.gov/nationsreportcard/subject/field_pubs/sqb/pdf/2014_sq_cghs_g8.pdf

Casner-Lotto, J., & Barrington, L. (2006). *Are they really ready to work? Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century U.S. workforce*. Retrieved from http://www.conference-board.org/pdf_free/BED-06-Workforce.pdf

Chan, D. W. (2000). Assessing leadership among chinese secondary students in Hong Kong: The use of the roets rating scale for leadership. *Gifted Child Quarterly, 44*(2), 115–122.

Ciancialo, A., & Sternberg, R. J. (2004). *Intelligence: A brief history*. Boston: Wiley-Blackwell.

Common Core State Standards Initiative (2010a). *Common core state standards for english language arts & literacy in history/social studies, science, and technology subjects*. Downloaded November 30, 2014, from http://www.corestandards.org/wp-content/uploads/ELA_Standards.pdf

Common Core State Standards Initiative (2010b). *Common core state standards for mathematics*. Downloaded November 30, 2014, from http://www.corestandards.org/wp-content/uploads/Math_Standards.pdf

Covay Minor, E., Porter, A. C., Murphy, J., Goldring, E. B., Cravens, X., & Elloitt, S. N. (2014). A known group analysis validity study of the vanderbilt assessment of leadership in education in US elementary and secondary schools. *Educational Assessment, Evaluation and Accountability, 26*(1), 29–48.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. JSAS Catalog of Selected Documents in Psychology, 1980, 10, 85.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113–126.

de Wied, M., Maas, C., van Goozen, S., Vermande, M., Engels, R., Meeus, W., Matthys, W., & Goudena, P. (2012). Bryant's empathy index: A closer examination of its internal structure. *European Journal of Psychological Assessment, 23*(2), 99–104. Downloaded August 10, 2015, from http://devpsychopathologyru.nl/wp-content/uploads/2012/11/2007-Bryant%E2%80%99s-Empathy-Index.pdf

deMarrais, K. B., & LeCompte, M. D. (1995). *The way schools work: A sociological analysis of education* (2nd ed.). White Plains, NY: Longman Publishers.

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., et al. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *Journal of Autism Developmental Disorder, 38*(3), 464–473.

Elliott, A. J., & Sheldon, K. M. (1997). Avoidance achievement motivation: A personal goals analysis. *Journal of Personality and Social Psychology, 73*(1), 171–185.

Gardner, P. D. (2007). *Recruiting trends: 2006–2007*. East Lansing, MI: Collegiate Employment Research Institute, Michigan State University.

Gibbs, J. C., Clark, P. M., Joseph, J. A., Green, J. L., Goodrick, T. S., & Makowski, D. G. (1986). Relations between moral judgment, moral courage, and filed independence. *Child Development, 57*, 1040–1043.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.

Goals 2000: Educate America Act, H.R. 1804 (1994).

Goodlad, J. I. (1979). *What schools are for*. Bloomington, IN: Phi Delta Kappa Educational Foundation.

Guglielmino, L. M. (1978). Development of the self-directed learning readiness scale. (Doctoral dissertation, University of Georgia, 1977). *Dissertation Abstracts International, 38,* 6467A.

Hills, P. R., Francis, L. J., & Jennings, P. (2011). The school short-form coopersmith self-esteem inventory: Revised and improved. *Canadian Journal of School Psychology, 26*(1), 62–71.

Howe, A., Pitten Cate, I. M., Brown, A., & Hadwin, J. A. (2008). Empathy in preschool children: The development of the southampton test of empathy for preschoolers (STEP). *Psychological Assessment, 20*(3), 305–309.

Immerwahl, J. (2000). *Great expectations: How Californians view higher education*. San Jose, CA: National Center for Public Policy and Higher Education and Public Agenda.

International Association for the Evaluation of Educational Achievement. (1999). *IEA Civic education study* (No. ICPSR 3892). Ann Arbor, MI: Interuniversity consortium for political and social research.

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the basic empathy scale. *Journal of Adolescence, 29*, 589–611.

Kemler, E. R. (2015). *Defining and measuring responsible citizenship in higher education contexts in the US and Argentina*. Undergraduate thesis, Wesleyan University.

Klein, A. (2013, May 14). Sequestration forces cuts to national social studies tests. *Education Week*. Downloaded May 28, 2015, from http://blogs.edweek.org/edweek/campaign-k-12/2013/05/social_studies_naep_tests_post.html

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. Goslin (Ed.), *Handbook of socialization theory and research*. New York: Rand McNally.

Kurland, H., Peretz, H., & Hertz-Lazarowtiz, R. (2010). Leadership style and organizational learning: The mediate effect of school vision. *Journal of Educational Administration, 48*(1), 7–30.

Labaree, D. F. (1997). *How to succeed in school without really learning*. New Haven, CT: Yale University Press.

Leontopoulou, S. (2010). An exploratory study of altruism in greek children: Relations with empathy, resilience and classroom climate. *Psychology, 1*(5), 377–385.

Logan, S., Medford, E., & Hughes, N. (2011). The importance of intrinsic motivation for high and low ability readers' reading comprehension performance. *Learning and Individual Differences, 21*(1), 124–128.

Lounsbury, J. W., Saudargas, R. A., Gibson, L. W., & Leong, F. T. (2005). An investigation of broad and narrow personality traits in relation to general and domain-specific life satisfaction of college students. *Research in Higher Education, 46*(6), 707–729.

Lounsbury, J. W., Levy, J. J., Park, S. H., Gibson, L. W., & Smith, R. (2009). An investigation of the construct validity of the personality trait of self-directed learning. Learning and Individual Differences, 19(4), 411–418. doi:10.1016/j.lindif.2009.03.001.

MacCann, C., Fogarty, G. J., Zeidner, M., & Roberts, R. (2011). How emotional intelligence and coping affect academic success. *Contemporary Educational Psychology, 36*, 60–70.

Mavroveli, S., Petrides, K. V., Sangareau, Y., & Furnham, A. (2009). Exploring the relationships between trait emotional intelligence and objective socio-emotional outcomes in childhood. *British Journal of Educational Psychology, 79*(2), 259–272.

Mavroveli, S., & Sanchez-Ruiz, M. J. (2011). Trait emotional intelligence influences on academic achievement and school behaviour. *British Journal of Educational Psychology, 81*, 112–134.

Mayer, J., Salovey, P., & Caruso, D. (2002). *Mayer-Salovey-Caruso emotional intelligence test manual*. Toronto, ON: Multi-Health Systems.

McAuley, E., Duncan, T., & Tammen, V. V. (1987). Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport, 60*, 48–58.

McClough, A. C., & Rogelberg, S. G. (2003). Selection in teams: An exploration of the teamwork knowledge, skills, and ability test. *International Journal of Selection and Assessment, 11*(1), 56–66.

McIntosh, S. (2012). *State high school exit exams: A policy in transition*. Washington, DC: Center on Educational Policy.

Mehrabian, A. (1996). *Manual for the balanced emotional empathy scale (BEES)*. Monterey, CA: Albert Mehrabian.

Mestre, V., Frias, M. D., & Samper, P. (2004). La medida de la empatia: Analysis del interpersonal reactivity index [Measuring empathy: Analysis of the interpersonal reactivity index]. *Psicothema, 16*, 255–260.

National Commission on Excellence in Education. (1983). A Nation at Risk: The Imperative for Educational Reform. Washington, DC: US Department of Education.

Noddings, N. (1995). *Philosophy of education*. Boulder, CO: Westview Press.

Ojanen, T., Sijtsema, J. J., Hawley, P. H., & Little, T. D. (2010). Intrinsic and extrinsic motivation in early adolescents' friendship development: Friendship selection, influence, and prospective friendship quality. *Journal of Adolescence, 33*(6), 837–851.

Parker, J. D. A., Creque, R. E., Barnhart, D. L., Harris, J. I., Majeski, S. A., Wood, L. M., et al. (2004). Academic achievement in high school: Does emotional intelligence matter? *Personality and Individual Differences, 37*(7), 1321–1330.

PISA (2012). *PISA 2012 Results: What students know and can do*. Paris: OECD.

Plenty, S., & Heubeck, B. G. (2011). Mathematics motivation and engagement: An independent evaluation of a complex model with Australian rural high school students. *Educational Research and Evaluation, 17*(4), 283–299.

Rae Westbury, H., & Neumann, D. L. (2008). Empathy-related responses to moving film stimuli depicting human and non-human animal targets in negative circumstances. *Biological Psychology, 78*(1), 66–74.

Rameson, L. T., Morelli, S. A., & Lieberman, M. D. (2012). The neural correlates of empathy: Experience, automaticity, and prosocial behavior. *Journal of Cognitive Neuroscience, 24*(1), 235–245.

Reed, R. F., & Johnson, T. W. (Eds.). (1996). *Philosophical documents in education*. White Plains, NY: Longman Publishers.
Renchler, R. (1991). Leadership with a vision: How principals develop and implement their visions for school success. *OSSC Bulletin, 34*(5), 1–29.
Renihan, P. J., Renihan, F. I., & Waldron, P. (1986). The common ingredients of successful school effectiveness projects. *Education Canada, 26*(3), 16–21.
Rest, J. (1979). *Development in judging moral issues*. Minneapolis, MN: University of Minnesota Press.
Rest, J., Narvaez, D., Bebeau, M., & Thoma, S. (1999). DIT-2: Devising and testing a new instrument of moral judgment. *Journal of Educational Psychology, 91*(4), 644–659.
Robinson, J.P., Shaver, P.R., & Wrightsman, L.S. (1991). (Eds). Measures of Personality and Social Psychological Attitudes. San Diego: Elsevier.
Roets, L. (1986). *Roets rating scale for leadership*. Des Moines, IA: Leadership Publishers.
Rojas, R., & Rich, M. (2015, January 27). States move to make citizenship exams a classroom aid. *New York Times*. Accessed May 28, 2015, from http://www.nytimes.com/2015/01/28/us/states-move-to-make-citizenship-exams-a-classroom-aid.html
Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
Rousseau, C., Drapeau, A., Lacroix, L., Bagilishya, D., & Heusch, N. (2005). Evaluation of a classroom program of creative expression workshops for refugee and immigrant children. *Journal of Child Psychology and Psychiatry, 46*(2), 180–185.
Rutter, M., & Maughan, B. (2002). School effectiveness findings 1979–2002. *Journal of School Psychology, 40*(6), 451–475.
Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology, 57*(5), 749–761.
Ryan, R. M., Mims, B., & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology, 45*, 736–750.
Schmitt, N. (2012). Development of rationale and measures of noncognitive college student potential. *Educational Psychologist, 47*(1), 18–29.
Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1–65). New York: Academic.
Schwartz, S. (2007). Value priorties and behavior: Applying a theory of integrated value systems. In C. Seligman, J.M. Olson, and M.P. Zanna (2013). The Psychology of Values: The Ontario Symposium, Vol. 8, pp. 1–24.
Schwartz, S., & Bardi, A. (2001). Value hierarchies across cultures: Taking a similarities perspective. *Journal of Cross-Cultural Psychology, 32*(3), 268–290.
Sebastian, J., & Allensworth, E. (2012). The influence of principal leadership on classroom instruction and student learning: A study of mediated pathways to learning. *Educational Administration Quarterly, 48*(4), 626–663.
Shrauger, J., & Schohm, M. (1995). Self-confidence in college students: Conceptualization, measurement, and behavioral implications. *Assessment, 2*(3), 255–278.
Sperling, R. A., Sherwood, T. P., & Hood, A. M. (2013). Relating motivation to read and emergent reading skills: A measurement validation study. *Reading Psychology, 34*(5), 461–485.
Stemler, S. E. (2001). Examining school effectiveness at the fourth grade: A hierarchical analysis of the third international mathematics and science study (TIMSS). (Doctoral dissertation, Boston College, 2001). *Dissertation Abstracts International, 62(03A), 919*.
Stemler, S. E. (2012a). What should university admissions tests predict? *Educational Psychologist, 47*(1), 5–17.
Stemler, S. E. (2012b). *Using situational judgment testing to measure ethical values among college students.* Paper presented at the annual meeting of the National Council on Measurement in Education (NCME), New Orleans, LA.

Stemler, S. E. (2015). *Measuring creativity and citizenship in K-12 students*. Final report of the evaluation of the Higher Order Thinking Skills (HOTS) schools in connecticut.

Stemler, S. E., & Bebell, D. (1999, April). *An empirical approach to understanding and analyzing the mission statements of selected educational institutions*. Paper presented at the New England Educational Research Organization (NEERO), Portsmouth, NH.

Stemler, S. E., & Bebell, D. (2012). *The school mission statement: Values, goals, and identities in American education*. New York: Routledge.

Stemler, S. E., Bebell, D., & Sonnabend, L. (2011). Using mission statements for reflection and research. *Educational Administration Quarterly, 47*(2), 383–420.

Stevens, M. J., & Campion, M. A. (1999). Staffing work teams: Development and validation of a selection test for teamwork settings. *Journal of Management, 25*, 207–228.

Stober, S. S. (1997). *A content analysis of college and university mission statements*. Unpublished doctoral dissertation, Temple University, Philadelphia, PA.

Straka, G. A. (1995). *Problems of measuring self-directed learning readiness*. Paper presented at the Asia-Pacific seminar on self-directed learning, Seoul, Korea.

Teddlie, C., & Reynolds, D. (Eds.). (2000). *The international handbook of school effectiveness research*. New York: Falmer Press.

Tejeda, M. J., Scandura, T. A., & Pillai, R. (2001). The MLQ revisited: Psychometric properties and recommendations. *The Leadership Quarterly, 12*, 31–52.

Tietjen, A. M., & Walker, L. J. (1985). Moral reasoning and leadership among men in a Papua New Guinea society. *Developmental Psychology, 21*, 982–989.

Timothy, T., Chee, T. S., Beng, L. C., Sing, C. C., Ling, K. J. H., Li, C. W., et al. (2010). The self-directed learning with technology scale (SDLTS) for young students: An initial development and validation. *Computers & Education, 55*(4), 1764–1771.

Tirri, K., & Nokelainen, P. (2007). Comparison of academically average and gifted students' self-rated ethical sensitivity. *Educational Research and Evaluation, 13*(6), 587–601.

Tolan, P. H., & Larsen, R. (2014). Trajectories of life satisfaction during middle school: Relations to developmental-ecological microsystems and student functioning. *Journal of Research on Adolescence, 24*(3), 497–511.

Tyack, D. B. (1988). Ways of seeing: An essay on the history of compulsory schooling. In R. M. Jaeger (Ed.), *Complementary methods for research in education* (pp. 24–59). Washington, DC: American Educational Research Association.

United States Citizenship and Immigration Services. (2014). *Civics test.* Retrieved November 25, 2014, from http://www.uscis.gov

Wang, L., MacCann, C., Zhuang, X., Liu, O. L., & Roberts, R. D. (2009). Assessing teamwork and collaboration in high school students: A multimethod approach. *Canadian Journal of School Psychology, 24*(2), 108–124.

West, R. F., & Bentley, E. I. (1990). Structural analysis of the self-directed learning readiness scale: A confirmatory factor analysis using LISREL modeling. In H. B. Long (Ed.), *Advances in research and practice in self-directed learning* (pp. 157–180). Norman, OK: Oklahoma Research Center for Continuing Professional and Higher Education of the University of Oklahoma.

Widmer, M. A., Ellis, G. D., & Munson, W. W. (2003). Development of the Aristotelian ethical behavior in leisure scale short form. *Therapeutic Recreation Journal, 37*(3), 256–274.

Widmer, M. A., Ellis, G. D., & Trunnell, E. P. (1996). Measurement of ethical behavior in leisure among high- and low-risk adolescents. *Adolescence, 31*(122), 397–408.

Woods, S., Wolke, D., Nowicki, S., & Hall, L. (2009). Emotion recognition abilities and empathy of victims of bullying. *Child Abuse and Neglect, 33*(5), 307–311.

Xu, Y., Iran-Nejad, A., & Thoma, S. J. (2007). Administering defining issues test online: Do response modes matter? *Journal of Interactive Online Learning, 6*(1), 10–27.

Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M., Brandon, P. R., Furtak, E. M., et al. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education, 21*(4), 335–359.

Zeidner, M., Matthews, G., & Roberts, R. D. (2009). *What we know about emotional intelligence: How it affects learning, work, relationships, and our mental health*. Cambridge, MA: MIT Press.

Chapter 4
Personality Development During the School-Aged Years: Implications for Theory, Research, and Practice

Kate E. Walton and Kimberly A. Billera

4.1 A Developmental Trait Taxonomy

Historically, childhood temperament was studied in isolation from adult personality. Temperament has been generally defined as biologically determined traits, which are present in infancy and show temporal and situational consistency. Temperamental traits were most often studied by child or developmental psychologists. Personality has been defined as "the dynamic organization within the individual of those psychophysical systems that determine his characteristic behavior and thought" (Allport, 1961, p. 28), and personality psychologists typically study traits in adulthood. More recent advances have lead to the synthesis of child temperament and adult personality research. A key finding relevant for this particular chapter concerns the hierarchical nature of childhood temperament traits and how this hierarchy maps onto adult personality traits. The similarity in structure across the lifespan can facilitate our discussion of the development of personality.

Within the past few decades, the field has reached consensus concerning the structure of adult personality traits. Multiple personality trait taxonomies have been put forth with the number of key personality traits ranging from as few as one (Musek, 2007) or two (Block & Block, 1980b; Digman, 1997) to as many as 20 (Gough, 1987). These seemingly distinct taxonomies can be reconciled with the understanding that personality is structured hierarchically, much like intelligence. There is thought to be a general factor of intelligence (*g*; Spearman, 1904), which subsumes more narrow and specific factors. For example, Carroll (1993) posits a three-stratum model of cognitive ability; 69 narrow abilities make up stratum I, eight broad abilities make up stratum II, and *g* lies in stratum III. For example, fluid

K.E. Walton (✉) • K.A. Billera
Department of Psychology, St. John's University,
8000 Utopia Parkway, Jamaica, NY 11439, USA
e-mail: waltonk@stjohns.edu; kimberlybillera@gmail.com

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_4

reasoning is one of the eight broad abilities, and it comprises the three narrow abilities of induction, general sequential reasoning, and quantitative reasoning. Personality is similarly structured. At the most general level of the hierarchy is a general personality factor, referred to as the Big One (Musek, 2007). Beneath the Big One lie two factors often referred to as alpha and beta (Digman, 1997). Alpha and beta are thought to be associated with socialization processes and personal growth, respectively. At the next level of the hierarchy, the more narrow traits of agreeableness, conscientiousness, and emotional stability load on the alpha factor, while the more narrow traits of extraversion and openness to experience load on beta. Some models of personality focus on a much greater number of specific traits, such as Cattell's 16-factor taxonomy (Cattell, 1979).

While broad traits have the advantage of summarizing large amounts of information, narrow traits allow for more nuanced description. This trade-off has long been recognized as the bandwidth-fidelity dilemma (Cronbach & Gleser, 1965). At a specific level of the hierarchy, traits reach an optimal balance of high bandwidth and high fidelity. This level has been termed the foundational level of personality structure (Soto & John, 2014). In the adult literature, the Big Five personality traits – extraversion, agreeableness, conscientiousness, emotional stability (with its opposite pole, neuroticism, most often referenced in the literature), and openness to experience – have been recognized as achieving this optimal balance, and consequently, this five-factor taxonomy is the most widely used in personality research (see Chap. 1). In addition to maximizing the bandwidth-fidelity balance, the Big Five have garnered tremendous empirical support in the past few decades, as the same five-factor structure is generalizable to both men and women and across numerous populations and settings (Marsh, Nagengast, & Morin, 2013; McCrae & Costa, 2003; McCrae, Terracciano, & 78 members of the Personality Profiles of Cultures Project, 2005).

While this five-factor model is well established in the adult literature and has facilitated studies of adult personality development, efforts to discern the foundational level of child and adolescent personality structure are more recent. Findings suggest that, like adult personality, childhood personality is hierarchical (Soto & John, 2014; Tackett, Krueger, Iacono, & McGue, 2008; Tackett et al., 2012). In studies of childhood temperament and personality, five factors emerge that are markedly similar to the Big Five personality traits identified in the adult literature. Adult extraversion stems from infant/childhood positive emotionality and surgency, which subsume the more specific facets of sociability and activity level. Adult agreeableness stems from childhood benevolence and need for affiliation, and adult conscientiousness develops from childhood tendencies to demonstrate constraint and effortful control. Infant and childhood negative emotionality, like adult neuroticism, incorporates feelings of fearfulness, anxiety, and sadness. Finally, childhood creativity and imagination pave the way for adult openness to experience (Caspi, Roberts, & Shiner, 2005; Caspi & Shiner, 2006; De Pauw & Mervielde, 2010).

Our goal in this chapter is to provide an overview of findings concerning the development of personality across the life course. Given the overlap from childhood to adulthood, the Big Five provide a feasible framework for this discussion. In

the following sections, we consider literature pertaining to the continuity and change of extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience (and related constructs) across the lifespan. It is worthwhile to note that in all ensuing discussion and review of the literature, we refer to emotional stability and reverse any interpretations that pertain to neuroticism.

4.2 Personality Continuity and Change

There is an array of methods to evaluate personality continuity and change, and it is critical to consider multiple methods to gain a full appreciation of the nature of personality development. To illustrate the need for considering different types of continuity and change, consider the trait of conscientiousness. Conscientiousness is defined as "socially prescribed impulse control that facilitates task- and goal-directed behavior, such as thinking before acting, delaying gratification, following norms and rules, and planning, organizing, and prioritizing tasks" (John & Srivastava, 1999, p. 121). It would not be a stretch to imagine that, as children age, they will show a greater tendency to follow rules, delay gratification, and consider the consequences of an action before carrying it out. That is, it is likely that one would observe an absolute increase in conscientiousness during the school-aged years. However, assume that across this time period, a child we will call Ben maintains his position of being lowest among his peers in terms of conscientiousness as he is consistently the most impulsive, disorganized, etc. That is, relative to his peers, he maintains a low level of conscientiousness. Therefore, in addition to the absolute change, it is possible to simultaneously observe relative stability. This distinction is necessary. Only considering absolute change would lead to the conclusion that conscientiousness is dynamic, changing across time. In contrast, only considering Ben's relative standing among his peers would lead to the conclusion that conscientiousness is static, unchanging across time. In the ensuing sections, we review findings from both perspectives and consider two additional methods of personality development assessment.

4.2.1 Rank-Order Stability

The above example of Ben maintaining his position of being lowest among his peers in conscientiousness illustrates rank-order stability. Rank-order stability reflects the degree to which the relative order of individuals on a given trait is maintained across time. Correlation coefficients are typically used as an index of personality continuity with high test-retest correlations indicating high rank-order stability or personality consistency across time.

Hampson and Goldberg (2006) gathered teachers' assessments of personality in first and second grade children, and over a 4-year period, stability coefficients

Table 4.1 Rank-order stability estimates given as test-retest correlations

Reference	Mean age	Interval length	Sex	E	A	C	ES	O
Hampson and Goldberg (2006)	6.5[a]	1	M&F	0.52	0.51	0.53	0.28	0.49
	6.5[a]	4	M&F	0.36	0.45	0.41	0.38	0.55
	7.0[a]	3	M&F	0.38	0.43	0.31	0.22	0.33
Prinzie and Deković (2008)	7.0	3	M&F	0.49	0.42	0.54	0.38	0.59
McCrae et al. (2002)	12.0	4	M	0.39	0.31	0.49	0.36	0.45
	12.0	4	F	0.45	0.34	0.63	0.30	0.34
Pullmann et al. (2006)	12.4	2	M&F[b]	0.49	0.51	0.48	0.50	0.57
	14.4	2	M&F[b]	0.62	0.49	0.59	0.64	0.48
	16.1	2	M&F[b]	0.73	0.62	0.69	0.65	0.68
Klimstra et al. (2009)	12.4	1	M	0.35	0.31	0.45	0.32	0.36
	13.4[c]	1	M	0.48	0.48	0.59	0.39	0.51
	14.4[c]	1	M	0.62	0.53	0.69	0.50	0.61
	15.4[c]	1	M	0.60	0.47	0.64	0.51	0.61
	12.4	1	F	0.55	0.41	0.55	0.48	0.52
	13.4[c]	1	F	0.67	0.46	0.60	0.59	0.54
	14.4[c]	1	F	0.68	0.51	0.67	0.56	0.60
	15.4[c]	1	F	0.75	0.53	0.72	0.66	0.69
Wortman et al. (2012)[d]	15–19[e]	4	M&F	0.70	0.61	0.58	0.58	0.70

Note: Mean age is at time 1 and is given in years. Interval length is given in years
M males, *F* females, *E* extraversion, *A* agreeableness, *C* conscientiousness, *ES* emotional stability, *O* openness
[a]Mean age not explicitly stated but inferred from year in school
[b]Males' and females' coefficients also presented separately in original paper
[c]Mean age not explicitly stated but inferred from interval length
[d]Older cohorts also presented in original paper
[e]Only age range presented in original paper

ranged from 0.36 for extraversion to 0.55 for openness. In a separate 3-year longitudinal study of continuity and change in children, teacher ratings of the Hierarchical Personality Inventory for Children (HiPIC; Mervielde & De Fruyt, 1999) were collected on children who were aged 6–9 years during the first wave of assessment (Prinzie & Deković, 2008). Three-year test-retest correlations ranged from 0.38 for emotional stability to 0.59 for imagination, a trait that maps onto openness (De Fruyt et al., 2006; Prinzie & Deković, 2008). These stability coefficients and those reported in the studies reviewed below are presented in Table 4.1.

In an older, adolescent sample of Estonian students, Pullmann, Raudsepp, and Allik (2006) studied the personality development of three groups (12-, 14-, and 16-year-olds) across a 2-year span. For the combined male and female sample, the average test-retest correlations of scores on the Estonian NEO-Five-Factor Inventory (NEO-FFI; Allik, Laidra, Realo, & Pullmann, 2004; Costa & McCrae,

1992) ranged from 0.48 (conscientiousness, from age 12 to 14) to 0.73 (extraversion, from age 16 to 18). In a second study carried out over a 5-year span during roughly the same adolescent time period, two cohorts were observed five times from early (12.4 years, on average) to middle adolescence and from middle (16.7 years, on average) to late adolescence (Klimstra, Hale III, Raaijmakers, Branje, & Meeus, 2009). Test-retest correlations of scores on the Dutch version of Goldberg's Big Five questionnaire (Gerris et al., 1998; Goldberg, 1992) for girls in the younger cohort ranged from 0.41 (Time 1–2, agreeableness) to 0.75 (Time 4–5, extraversion), and they ranged from 0.52 (Times 1–2 and 3–4, agreeableness) to 0.86 (Time 4–5, conscientiousness) for the older cohort. For boys, test-retest correlations in the younger cohort ranged from 0.31 (Time 1–2, agreeableness) to 0.69 (Time 3–4, conscientiousness), and they ranged from 0.27 (Time 1–2, agreeableness) to 0.75 (Time 3–4, conscientiousness, and Time 4–5, extraversion) for the older cohort. Similar estimates were derived in a sample of children with an average age of 12 years who were reassessed 4 years later with the NEO-FFI (McCrae et al., 2002); for girls, estimates ranged from 0.30 for emotional stability to 0.63 for conscientiousness, and for boys, estimates ranged from 0.31 for agreeableness to 0.49 for conscientiousness.

Wortman, Lucas, and Donnellan (2012) employed a sample with a much wider age range and assessed personality traits twice over a 4-year span with a variant of Saucier's (1994) adjective-based Big Five measure. Individuals ranged from age 15 to 84 years and were divided into 14 groups. The 15–19-year-old group exhibited a great deal of stability. Their average test-retest correlations reached 0.58 for conscientiousness and emotional stability, 0.61 for agreeableness, and 0.70 for extraversion and openness. Consistency estimates increased for the subsequent age groups, though these age-related changes in personality stability ultimately showed clear curvilinear trends for all of the Big Five traits. Stability estimates increased after the 15–19-year-old group, became larger during midlife, and then declined for the oldest cohort (80–84-year-olds). Peak stability generally emerged around the sixth decade of life when estimates reached 0.90 for extraversion, 0.85 for conscientiousness, 0.82 for openness, and 0.75 for both agreeableness and emotional stability.

To summarize the findings from multiple studies of personality consistency across time, an extensive meta-analysis was carried out in which the authors compiled more than 3000 test-retest correlations from 152 longitudinal studies (Roberts & DelVecchio, 2000). The average test-retest correlation (after controlling for length of interval) was 0.31 for the 0–2.9-year age range, 0.49 for the 3–5.9-year age range, and 0.43 for the 6–11.9 and 12–17.9-year age ranges. Thereafter, stability estimates increased until the sixth decade when they peaked at 0.74 and then showed a slight decline. Across the age groups, extraversion showed the greatest stability with an average test-retest correlation of 0.55, and emotional stability showed the lowest with an average of 0.46. Consistency estimates were negatively related to interval length, but there were no moderating effects for attrition, gender, or method of assessment.

This meta-analysis and collection of empirical studies illustrates several key features regarding the rank-order stability of personality. First, we see that personality is fairly stable over time. Even in children as young as 6 years old, we observe stability estimates in the range of 0.38–0.59 (Prinzie & Deković, 2008). Second, for the most part, we see that personality attributes increase in stability with age, supporting what has been referred to as the cumulative continuity principle (Caspi et al., 2005). In studies of school-aged children, stability estimates for older cohorts are greater than those for younger cohorts (Klimstra et al., 2009; Pullmann et al., 2006). For example, Pullmann and colleagues (2006) reported that, across traits, the average test-retest correlation increased from 0.51 from age 12 to 14 to 0.56 from age 14 to 16 and reached 0.67 from age 16 to 18 years. While this age-related increase in stability is well documented, studies that extend beyond adolescence into middle and late adulthood suggest that this increase is nonlinear. Wortman and colleagues (2012) documented a decline in stability coefficients during the later stages of life, and Roberts' and DelVecchio's (2000) meta-analysis revealed a slight drop after the 50–59 decade. It should be noted, though, that not all studies offer support for the cumulative continuity principle. For example, with their study of personality continuity in childhood and adolescence, De Fruyt and colleagues (2006) provided evidence for age-related linear decreases in continuity or more complex cubic age-stability relationships.

In addition, it is worth highlighting findings concerning gender differences in patterns of rank-order stability. Roberts and DelVecchio (2000) reported that there were no gender differences in personality stability. However, in some individual studies focusing on school-aged children and adolescents, some gender differences have been observed. Several studies show that girls tend to be more stable than boys during this period of development (Klimstra et al., 2009; Pullmann et al., 2006). One study suggests the gender effect is trait-specific with girls exhibiting greater stability on some traits (extraversion, agreeableness, and conscientiousness) and boys exhibiting greater stability on others (emotional stability and openness; McCrae et al., 2002). Moreover, there is some evidence that age-related increases in rank-order stability seem to be stronger in girls than in boys (Klimstra et al., 2009; Pullmann et al., 2006). Greater stability for females during the school-aged years may be due to females' advantage with physical and neural development; that is, females' bodies and brains mature earlier than males'. Some psychological mechanisms might be at play as well. Klimstra and colleagues (2009) cited evidence (Burwell & Shirk, 2007) that adolescent girls have a greater tendency for self-reflection and suggested that repeated reflection on oneself could possibly lead to greater personality consistency.

Finally, in the empirical studies of school-aged children and adolescents reviewed here, there is no overwhelmingly clear pattern concerning which traits show the greatest amount of consistency. Roberts and DelVecchio (2000) concluded that extraversion and agreeableness are the most consistent but that the remaining Big Five are also quite consistent with comparable stability estimates.

4.2.2 Mean-Level Change

In the preceding section, we established that there is a considerable amount of rank-order stability across time. As previously noted, this does not preclude the possibility of significant mean-level change. Differences in the average value of a group across time can still be observed. Mean-level change is commonly assessed with *t*-tests and/or indexed with a standardized effect size, such as Cohen's *d*. Cohen's *d* indicates how many standard deviations two values differ from one another. In a study of mean-level change, a *d* of −0.50 would indicate a decrease of one-half a standard deviation from Time 1 to Time 2. Often these are derived from longitudinal studies in which the same group of individuals is tracked over time, though cross-sectional differences between age groups at a single time point may also be considered. All studies reviewed here are longitudinal unless otherwise stated.

Prinzie and Deković (2008) gave an account of the mean-level change for their sample of elementary school-aged children (6–9 years) who were assessed with teachers' HiPIC ratings twice in a 3-year period. They reported no significant change on conscientiousness or emotional stability but significant decreases on extraversion ($d = -0.17$) and imagination (akin to openness; $d = -0.15$) and a small yet significant increase on benevolence (akin to agreeableness; $d = 0.09$). These effect sizes and those reported in one other paper reviewed are presented in Table 4.2 (many of the other studies we review here did not report effect sizes).

There are multiple studies describing mean-level change during the adolescent years. Conscientiousness seems to show no change (De Fruyt et al., 2006; Klimstra et al., 2009; Pullmann et al., 2006) or a slight decrease (De Fruyt et al., 2006; McCrae et al., 2002), while extraversion seems to show no change (De Fruyt et al., 2006; McCrae et al., 2002) or a slight increase (Klimstra et al., 2009; Pullmann et al., 2006). Emotional stability findings mimic those of extraversion with evidence of stability (De Fruyt et al., 2006; McCrae et al., 2002; Pullmann et al., 2006) or an increase (De Fruyt et al., 2006; Klimstra et al., 2009; Pullmann et al., 2006). Likewise, openness generally remains unchanged (De Fruyt et al., 2006) or increases

Table 4.2 Mean-level change estimates given as Cohen's d effect sizes

Reference	Mean age	Interval length	Sex	E	A	C	ES	O
Prinzie and Deković (2008)	7.0	3	M&F	−0.17	0.09	−0.08	0.07	−0.15
Pullmann et al. (2006)[a]	12.4	2	M&F	0.24	−0.15	0.06	0.15	0.10
	14.4	2	M&F	0.14	0.03	−0.03	0.13	0.25
	16.1	2	M&F	−0.02	0.04	0.08	0.04	0.27

Note: Mean age is at time 1 and is given in years. Interval length is given in years
M males, *F* females, *E* extraversion, *A* agreeableness, *C* conscientiousness, *ES* emotional stability, *O* openness

[a]Cohen's *d* reported for neuroticism so the signs were reversed to reflect change in emotional stability

with age (Klimstra et al., 2009; McCrae et al., 2002; Pullmann et al., 2006). The findings for agreeableness are more mixed. Pullmann and colleagues (2006) reported a decrease over the 2-year period for their youngest cohort (the 12-year-olds) but no change for the two older cohorts (the 14- and 16-year-olds). De Fruyt and colleagues (2006) and McCrae and colleagues (2002) found no evidence of change in agreeableness, but Klimstra and colleagues (2009) wrote that their sample became much more agreeable from early to mid-adolescence and from mid- to late adolescence.

Utilizing a different methodology for evaluating mean-level change, Branje, Lieshout, and Gerris (2007) fit growth curve models to self- and other reports of a Dutch adaptation of Goldberg's (1992) Big Five questionnaire from age 11 to 17 years using an accelerated longitudinal design. Emotional stability showed no change, and agreeableness showed a linear increase in female self-reports only. A linear increase in conscientiousness was also found for both female self- and observer reports. While self-reported openness increased linearly for girls, observer-reported openness decreased linearly for boys. Boys' self- and observer reports of extraversion revealed a linear decrease, while girls' observer reports revealed a linear increase. Girls' self-reports of extraversion indicated a curvilinear trend with an initial increase followed by a decrease.

Thus far, we have detailed personality change in childhood and adolescence. Two cross-sectional studies covering a wide range of ages and extending into old age allow us to garner an understanding of how personality continues to change beyond the adolescent years. As cited above, Wortman and colleagues (2012) assessed personality traits of individuals aged 15–84 years, who were divided into 14 age groups, twice over a 4-year span. In addition to looking at mean-level changes within each group over the 4-year period, they examined cross-sectional trends across the entire sample, spanning 69 years. The older participants were less extraverted and open than younger participants. Emotional stability increased linearly with age, and conscientiousness generally increased across the life course. Agreeableness showed a positive trend across early and middle adulthood, but there was a plateau after age 50, followed by a slight decrease after age 70. In a second cross-sectional study, Soto and colleagues (2011) gathered Big Five Inventory (BFI; John, Donahue, & Kentle, 1991) self-reports from a large (the sample size exceeded one million) internet sample who ranged in age from 10 to 65 years. Extraversion showed a modest negative trend from late childhood into adolescence. Thereafter, little to no change was observed through early adulthood and into middle age. Despite conscientiousness's negative trend from late childhood into adolescence, there was a pronounced positive trend from adolescence and through adulthood. Agreeableness and openness showed patterns highly similar to that of conscientiousness yet with less pronounced changes. The pattern of change for emotional stability differed across males and females. Males became slightly more emotionally stable from childhood through adulthood. Females also became more emotionally stable through adulthood, but this was preceded by a decrease during childhood.

We see that there are some mixed findings concerning the nature of personality change across the life course (e.g., there is evidence that agreeableness increases,

decreases, and remains unchanged during adolescence). In an effort to clarify this, Roberts, Walton, and Viechtbauer (2006) carried out a meta-analysis of 92 longitudinal studies documenting mean-level personality change. Two distinct facets of extraversion were considered because of suspected (and subsequently confirmed) differences in maturation. Social vitality showed a slight increase during the college years (ages 18–21.9 years; $d=0.06$) and decreases during the age 22–30 ($d=-0.16$) and 60–70 ($d=-0.16$) periods. Social dominance, in contrast, showed steady increases during adolescence ($d=0.20$), the college years ($d=0.41$), the 20s ($d=0.28$), and the 30s ($d=0.18$). Increases across much of the life course were also observed for conscientiousness and emotional stability. Effect sizes for conscientiousness reached 0.22, 0.26, 0.10, and 0.22 for the 20s, 30s, 40s, and 60s, respectively. For emotional stability, they reached 0.16, 0.12, 0.23, 0.26, and 0.06 for the adolescent years, college years, 20s, 30s, and 50s, respectively. Openness did not show a distinct pattern of development; there was a strong increase during the college years ($d=0.37$) and a moderate decrease during the period of 60–70 ($d=-0.19$). Agreeableness showed the least amount of systematic change with the only significant change occurring during the 50–60 period ($d=0.30$). Finally, we should note that there was no evidence that men and women change in distinct ways; there was no relationship between gender and mean-level change. There also was no moderating effect of attrition, yet there were some effects for interval length and cohort standing.

The overview of these studies leads us to draw at least two conclusions. First, personality continues to change in adulthood, even into late adulthood in some cases (Roberts et al., 2006). Although we see significant changes during childhood and adolescence (Branje et al., 2007; De Fruyt et al., 2006; Klimstra et al., 2009; McCrae et al., 2002; Prinzie & Deković, 2008; Pullmann et al., 2006), young adulthood is the period for the most significant changes, countering the idea that personality development is a phenomenon restricted to childhood and adolescence (Roberts et al., 2006). The second noteworthy conclusion we can draw is that, much like a fine wine, people get better with age. People exhibit increased conscientiousness, emotional stability, and social dominance across much of the life course. Meta-analytic findings point to a positive (though not statistically significant) trend for agreeableness as well (Roberts et al., 2006). These findings support the maturity principle, which states that people demonstrate an increased capability of being a productive and involved member of society, an increased tendency to be planful and decisive, and a greater propensity for being considerate and charitable (Caspi et al., 2005).

4.2.3 Individual-Level Change

While rank-order stability and mean-level change entail the entire population or study sample, individual-level change focuses on the increase or decrease in a particular trait demonstrated by a particular individual. Individual- and mean-level changes are independent of one another, and it is possible to observe one without

the other. For a simplistic example, if all females in the population were to increase on a trait in a given time period yet all males were to exhibit a decrease of equal magnitude, these individual changes would cancel one another out, and there would be no significant mean-level change. Individual-level change is commonly estimated with a difference score (ideally corrected for measurement error). A common index is the Reliable Change Index (RCI; Jacobson & Truax, 1991). The index is calculated by dividing the difference between Time 1 and Time 2 scores by the standard error of the difference between the two scores. RCI scores are normally distributed under the expectation that there is no change. Therefore, scores within ±1.96 are unlikely to occur without true change, and scores exceeding ±1.96 (i.e., the extreme 5 % of the distribution) are thought to represent true change. Although we believe the level of individual change is continuous and do not intend to imply some natural trichotomy of increasers, decreasers, and those who remain the same, we refer to individuals with scores in the extreme 5 % of the distribution as "increasers" or "decreasers" for the sake of simplicity.

Prinzie and Deković (2008) found high individual-level continuity among children ages 6–9 years old in their 3-year longitudinal study. The majority of children, approximately 78 %, showed no change at the domain level. No child in the sample changed on all of the traits, 17 % of the children changed on one trait, and only 5 % changed on two traits. More children exhibited reliable change on benevolence (7.2 %) than on any other trait. Benevolence had both more increasers (4.2 %) and decreasers (3.0 %) than expected by chance. Emotional stability had more increasers than expected by chance (4.4 %). Conscientiousness (3.8 %) and imagination (4.0 %) had more decreasers than expected by chance. Extraversion did as well with 3.2 % exhibiting a reliable decrease. Extraversion was the trait on which the fewest number of children exhibited reliable change with a total of 4.4 %. These percentages and those explicitly presented in the studies reviewed below are presented in Table 4.3.

Results from De Fruyt and colleagues' (2006) study of personality development (assessed with parent HiPIC ratings) in childhood and adolescence (aged 6–13 years and split into four age groups) also suggest high levels of individual-level continuity. The majority of participants were stable across all traits. Only a single individual showed change on four of the five traits, and none of the participants exhibited reliable change on all traits. Approximately 20 % exhibited change on one trait, while 5–10 % (varying across the different age groups) reliability changed on two traits. Similar findings of high individual-level continuity were reported by Pullmann and colleagues (2006) in their 2-year study of 12-, 14-, and 16-year-olds. On average, across the three groups and across the Big Five, more than 82 % of the adolescents showed no reliable change. However, there were more increasers and decreasers than expected by chance for all three groups on all traits. Averaged across the three groups, openness showed the most amount of change with 6.1 % decreasing and 13.0 % increasing. Agreeableness had the fewest amount of individual-level change with 83.7 % remaining the same across the 2-year period.

McCrae and colleagues (2002) reported much higher amounts of individual-level change in their longitudinal study tracking adolescents from age 12 to 16 years. Of

Table 4.3 Individual-level change estimates given as a percentage of participants with reliable change index scores exceeding −1.96 (decrease), exceeding +1.96 (increase), or falling within those bounds (stable)

Reference	Mean age	Interval length	Sex	E			A		
				% decrease	% stable	% increase	% decrease	% stable	% increase
Prinzie and Deković (2008)	7.0	3	M&F	3.2	95.6	1.2	3.0	92.8	4.2
McCrae et al. (2002)[a]	12.0	4	M&F	18.3	66.0	15.7	20.4	62.2	17.4
Pullmann et al. (2006)[a]	12.4	2	M&F	6.1	79.7	14.2	11.16	82.6	5.8
	14.4	2	M&F	6.8	80.6	12.6	8.3	81.0	10.7
	16.1	2	M&F	9.7	82.8	7.5	6.6	87.5	5.9

Reference	C			ES			O		
	% decrease	% stable	% increase	% decrease	% stable	% increase	% decrease	% Ss	% increase
Prinzie and Deković (2008)	3.8	94.9	1.3	1.7	93.9	4.4	4.0	94.3	1.7
McCrae et al. (2002)	22.6	62.6	14.8	23.5[a]	56.5	20.0	5.2	51.7	43.5
Pullmann et al. (2006)	8.4	82.4	9.2	8.9[a]	78.8	12.4	6.6	85.2	8.2
	8.4	81.4	10.2	5.8[a]	80.1	14.1	8.3	74.7	17.0
	5.5	86.2	8.3	4.8[a]	85.9	9.3	3.4	82.8	13.8

Note: Mean age is at time 1 and is given in years. Interval length is given in years

M = males, *F* = females, *E* = extraversion, *A* = agreeableness, *C* = conscientiousness, *ES* = emotional stability, *O* = openness

[a]Scored in direction of neuroticism so % increase and decrease were reversed

the five traits, extraversion showed the least amount of change with 66 % of the sample remaining stable. For most traits, there were fairly equivalent (within 3.5 %) numbers of increases and decreases, but there were considerably more decreases (22.6 %) than increases (14.8 %) for conscientiousness and far more increases (43.5 %) than decreases (5.2 %) for openness.

Collectively, the four sets of individual-level change results reviewed here suggest that the majority of children and adolescents exhibit no reliable change across these 2–4-year periods. McCrae and colleagues (2002) reported more change than the other authors, which could be the result of their longer test-retest interval (e.g., 4 years vs. the 2-year interval in the Pullmann et al., 2006 study). Although in all studies the majority remained the same, there were still significantly more children and adolescents exhibiting reliable change than one would expect by chance. Worth noting is that not all children show normative changes. For example, a meta-analysis of mean-level personality change (Roberts et al., 2006) documented a 0.20 standard deviation increase in social dominance during the adolescent years, yet McCrae and colleagues (2002) reported that 18.3 % of their sample showed evidence of a reliable decrease in extraversion. In addition, while there is a normative 0.16 standard deviation increase in emotional stability during the adolescent years (Roberts et al., 2006), 23.5 % of McCrae and colleagues' (2002) sample increased in neuroticism during this period. So while most people remain stable and those who do change show change in the normative direction, there are individuals who develop in a non-normative manner. This again illustrates the importance of considering multiple indices of personality change and continuity; these individual-level changes can be masked when focusing solely on group-level changes.

4.2.4 Ipsative Continuity

The prior three methods of investigating personality continuity or change concentrate on a single trait at a time (e.g., one would calculate separate test-retest correlations for conscientiousness, extraversion, etc.). Ipsative change refers to alterations in the configuration of variables within an individual across time and relies on some type of profile analysis. *Q* correlations are commonly used as an indicator of profile similarity. A ranked set of traits at Time 1 are correlated with a ranked set of traits at Time 2. For example, using the California Q-sort (Block, 1978; Block & Block, 1980a), raters consider 100 cards listing descriptive personality statements and arrange the cards in groups ranging from statements that are least descriptive of the target individual to most descriptive. If this is carried out at two time points, the groups of statements at Times 1 and 2 can be correlated with one another. Higher correlations indicate a greater degree of stability of trait configuration across time.

California Child (Block & Block, 1980a) and Adult (Block, 1978) Q-sort profile scores have been obtained in several studies, and the stability of those profiles over time have been investigated. In one study, from age 3–4 to age 7 years, median *q* correlations reached 0.52 for boys and girls (Ozer & Gjerde, 1989), slightly higher

than the median correlation observed in another study in which age 4- and 6-year profiles were correlated with one another (median correlation = 0.43; Asendorpf & van Aken, 1991). These correlations tend to increase with age. Asendorpf and van Aken (1991) reported median correlations of 0.47 and 0.61 (depending on the rater) between age 10-year and age 12-year profiles. Ozer and Gjerde (1989) noted that between ages 14 and 18 years, profile similarity increased to 0.71 for girls and 0.68 for boys, and Block (1971) reported that average *q* correlations between early and late adolescence exceed 0.70. While on average there is moderate-high profile similarity during childhood and adolescence, it should be noted that there is considerable variation across individuals in these *q* correlations, and negative correlations are often observed (Asendorpf & van Aken, 1991; Block, 1971; Ozer & Gjerde, 1989).

In addition to Q sets, ipsative continuity of the Big Five has been studied. As well as investigating rank-order stability and mean- and individual-level change, Prinzie and Deković (2008) reported on the ipsative continuity of teacher ratings across 3 years for their sample of children aged 6–9 years at Time 1. The average level of profile consistency across the 18 facet scales of the HiPIC ranged from –0.61 to 0.97, with a median of 0.56. More than 45 % of the children had high correlations falling between 0.50 and 0.80, while 21 % of the sample showed profile correlations below 0.30 %, and 7 % of the sample showed negative profile correlations from Time 1 to Time 2.

Like Prinzie and Deković (2008), De Fruyt and colleagues (2006) reported on ipsative continuity in addition to rank-order stability and mean- and individual-level change. *Q* correlations for the 18 HiPIC facets indicated a high level of ipsative continuity. The median correlations ranged from 0.81 to 0.85. In addition to *q* correlations, Cronbach and Gleser's (1953) three indices of profile variation were employed. These indices quantify variation in profiles' elevation (average level of scores), scatter (variability of scores), and shape (patterning of scores). D^2 is sensitive to all three and is a measure of the squared differences between traits at Time 1 and Time 2. D'^2 is sensitive to scatter and shape differences only and is a measure of the squared differences between Time 1 and Time 2 profiles after each has been centered around its mean. D''^2 is sensitive to shape differences only and quantifies the squared differences between profiles after each profile has been standardized. De Fruyt and colleagues' results indicated that a large percentage of individuals had stable trait profiles. For the most part, there was an increase in the number of significantly changed profiles moving from the youngest (age 6–7 years) to the oldest (age 12–13 years) group. The percentage of children with D^2 indices exceeding the cutoff value indicating a significant profile change was 9.1 % for the 6–7-year-old group, 14.8 % for the 8–9-year-old group, 14.9 % for the 10–11-year-old group, and 16.7 % for the 12–13-year-old group. The percentage of children exceeding the cutoff values for D'^2 and D''^2 was lower. Across all age groups, less than 10 % exhibited change in the shape of the profile; any changes were primarily changes in elevation and scatter.

Klimstra and colleagues (2009) reported on the 5-year profile similarity of their sample, and three findings are worth noting. Profile similarity increased with age, and these increases were more pronounced in the early mid-adolescence period than during the mid-late adolescence period. Furthermore, girls tended to exhibit greater

across-time profile similarity than boys. For the younger group of boys (12.4 years old at Time 1), *q* correlations ranged from 0.42 (Time 1–2) to 0.63 (Time 4–5), and they ranged from 0.63 (Time 1–2) to 0.76 (Time 4–5) for the older group of boys (16.7 years old at Time 1). For the younger girls, *q* correlations ranged from 0.55 (Time 1–2) to 0.74 (Time 4–5), and for the older girls, they ranged from 0.73 (Time 1–2) to 0.82 (Time 4–5). Ozer and Gjerde (1989) also documented increased profile similarity with age and a greater increase in earlier ages (e.g., the median *q* correlation for girls increased from 0.52 from age 3–4 to age 7 years to 0.63 from age 7 to age 11 years, but the median correlation between age 14 and 18 was only 0.02 higher than that between age 11 and 14, increasing from 0.69 to 0.71). Ozer and Gjerde did not report noticeably higher profile similarities for girls than boys.

4.3 Explanations for Developmental Patterns

In summary, we have reviewed literature showing that there is a moderate level of rank-order stability in childhood and stability estimates increase through adolescence and adulthood. For the most part, individuals maintain a stable profile of traits across time, and like rank-order stability, profile stability increases with age. Despite the relative consistency individuals maintain among their peers, and the ipsative continuity observed, there is significant individual- and mean-level change across the life course, even into late adulthood. What mechanisms or processes can account for these developmental patterns? To conclude, we will offer a brief overview of potential explanations of continuity and change.

Rank-order stability has been attributed to both biological and environmental causes. The Five Factor Theory (McCrae & Costa, 1999), for example, argues that traits are solely affected by biological factors. In part, this argument is based on behavior genetic findings confirming that genetic factors have a significant influence on personality trait differences (Bouchard & Loehlin, 2001). Heritability estimates for the Big Five are typically in the range of 0.50, and these estimates exceed those for shared environmental factors. While an innate, biological basis of traits may contribute to their across-time continuity, this does not rule out the impact of environmental processes that may also contribute to continuity. One such process is described by the corresponsive principle (Caspi et al., 2005), which states that life experiences tend to enhance preexisting traits that lead individuals to those experiences in the first place. Caspi and colleagues offered the dominance-leadership relationship as an example. Dominant individuals are more likely to take on leadership roles. Then, as a result of being in these leadership positions, they become even more dominant. Life experiences do not occur at random, creating extensive personality transformations. Instead, traits that are already prominent become even more ingrained due to the experiences they elicit. This is analogous to Bandura's (1978) social-cognitive theory of reciprocal determinism where behaviors, the environment, and cognitions and other internal events all reciprocally influence one another.

As we have seen, this level of continuity does not preclude significant change across the life course, and there are both biological and environmental explanations

for this change. As discussed above, the Five Factor Theory (McCrae & Costa, 1999) suggests that there is a biological basis for personality traits. In addition, this theory suggests that observed change is due to genetic predispositions to change in particular ways. That is, personality traits have intrinsic paths of development impervious to environmental effects (McCrae et al., 2000). In contrast, it has been argued that environmental demands influence change. At similar ages, people generally encounter specific life experiences which are accompanied by a set of demands at similar ages. For example, for most people, with early adulthood come one's first job, marriage, and parenthood. In the face of these experiences, one is forced to "grow up," become more socially invested, and become more conscientious and emotionally stable (Roberts et al., 2006; Roberts & Wood, 2006). The increased levels of conscientiousness, emotional stability, and social dominance during early adulthood support this assertion (Roberts et al., 2006). Childhood and adolescence bring their own set of demands, and Klimstra and colleagues (2009) speculated about the age-related demands and the resulting normative personality changes (or lack thereof). For instance, they reported no mean-level change in conscientiousness. For most adolescents, conscientiousness is most relevant for school work. The importance of academic achievement is constant across the entire adolescent period, which could explain why they and others found conscientiousness to be stable during this period. In contrast to some other researchers, Klimstra and colleagues (2009) found an increase in agreeableness during this period, and they argued that this is intuitive. During adolescence, peer relations become ever more salient and intimate, and agreeableness is key to establishing and maintaining positive, intimate relationships. The relationships and experiences of childhood and adolescence can at least partially dictate the nature of personality development.

4.4 Implications for the Academic Domain

The findings reviewed above have implications for many important outcomes, including academic achievement. Personality is known to be related to academic success and educational attainment. Poropat (2009) provided a meta-analysis of the relationship between the Big Five and academic performance. He concluded that academic performance is associated with agreeableness and openness but even more strongly with conscientiousness. Richardson, Abraham, and Bond (2012) carried out a second meta-analysis and concluded that conscientiousness is the only Big Five trait with significant correlations with university students' GPA. Both meta-analyses documented that the association between conscientiousness and academic performance is nearly as strong as that between intelligence and academic performance. Conscientiousness has also been shown to positively predict educational attainment (Lodi-Smith et al., 2010; Tross, Harper, Osher, & Kneidinger, 2000). The strength of association between conscientiousness and academic success remains fairly constant across academic levels, while associations between the other Big Five and academic performance tend to decrease from primary to secondary and tertiary school (Poropat, 2009).

Conscientiousness, the Big Five trait showing the strongest link with educational achievement and attainment, seems to show little to no change during the school-aged years (De Fruyt et al., 2006; Klimstra et al., 2009; Prinzie & Deković, 2008; Pullmann et al., 2006; Roberts et al., 2006). Roberts and colleagues (2006) reported population estimates of mean-level change in conscientiousness of 0.03 for the 10–18-year age range and 0.04 for the 18–22-year age range. Steady increases in conscientiousness are observed in the subsequent decades, though by this point, most individuals have completed their schooling. Therefore, in an effort to enhance students' scholastic performance, educators, school administrators, and school psychologists might consider personality-targeted interventions (see Chaps. 12 and 13). Attempts to expedite the maturation of conscientiousness, for example, may lead to greater academic success. A review of the literature has led us to the conclusion that personality traits are not set in stone but instead continue to change even into late adulthood. Not only do personality traits mature on their own, through no active interventions or programs designed to induce change, but research has shown that we can actively manipulate change in personality traits. Indeed, studies have shown that experimental (Jackson, Hill, Payne, Roberts, & Stine-Morrow, 2012), pharmacological (Tang et al., 2009), and therapeutic interventions (De Fruyt, van Leeuwen, Bagby, Rolland, & Rouillon, 2006) are effective in altering personality traits.

4.5 Conclusion

Using the five-factor trait taxonomy to guide our discussion, we have presented a depiction of personality development across the life course, highlighting the school-aged developmental period. During this period, personality traits show moderate-high levels of rank-order and profile stability. Reliable individual-level and significant mean-level change are also observed. Further demonstrating the independence of continuity and change, results signify that there is a greater degree of both stability and change in adulthood than in childhood and adolescence. This underscores the importance of employing multiple methods of investigating personality development. While personality evolves across the life course, its role in determining important life outcomes remains constant. To some degree, its evolution can be manipulated to enhance life outcomes, such as educational success.

References

Allik, J., Laidra, K., Realo, A., & Pullmann, H. (2004). Personality development from 12 to 18 years of age: Changes in mean levels and structure of traits. *European Journal of Personality, 18*, 445–462.

Allport, G. W. (1961). *Pattern and growth in personality*. Oxford, England: Hold, Reinhart, and Winston.

Asendorpf, J. B., & van Aken, M. A. G. (1991). Correlates of the temporal consistency of personality patterns in childhood. *Journal of Personality, 59*, 689–703.

Bandura, A. (1978). The self system in reciprocal determinism. *American Psychologist, 33*, 344–358.
Block, J. (1971). *Lives through time*. Berkeley, CA: Bancroft Books.
Block, J. (1978). *The Q-sort method in personality assessment and psychiatric research*. Palo Alto, CA: Consulting Psychologists Press.
Block, J., & Block, J. H. (1980a). *The California child Q-set*. Palo Alto, CA: Consulting Psychologists Press.
Block, J. H., & Block, J. (1980b). The role of ego-control and ego resiliency in the organization of behavior. In W. A. Collins (Ed.), *Minnesota symposia on child psychology* (pp. 39–101). Hillsdale, NJ: Erlbaum.
Bouchard, T. J., Jr., & Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Genetics, 31*, 243–273.
Branje, S. J., Van Lieshout, C. F., & Gerris, J. R. (2007). Big Five personality development in adolescence and adulthood. *European Journal of Personality, 21*, 45–62.
Burwell, R. A., & Shirk, S. R. (2007). Subtypes of rumination in adolescence: Associations between brooding, reflection, depressive symptoms, and coping. *Journal of Clinical Child and Adolescent Psychology, 36*, 56–65.
Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytical studies*. New York: Cambridge University Press.
Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology, 56*, 453–484.
Caspi, A., & Shiner, R. L. (2006). Personality development. In W. Damon, R. Lerner, & N. Eisenberg (Eds.), *Handbook of child psychology: Vol. 3 social, emotional, and personality development* (6th ed., pp. 300–364). New York: Wiley.
Cattell, R. B. (1979). *Cattell 16 personality factor questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.
Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
Cronbach, L. J., & Gleser, G. C. (1953). Assessing similarity between profiles. *Psychological Bulletin, 50*, 456–473.
Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions* (2nd ed.). Urbana, IL: University of Illinois Press.
De Fruyt, F., Bartels, M., Van Leeuwen, K. G., De Clercq, B., Decuyper, M., & Mervielde, I. (2006). Five types of personality continuity in childhood and adolescence. *Journal of Personality and Social Psychology, 91*, 538.
De Fruyt, F., Van Leeuwen, K., Bagby, R. M., Rolland, J.-P., & Rouillon, R. (2006). Assessing and interpreting personality change and continuity in patients treated for major depression. *Psychological Assessment, 18*, 71–80.
De Pauw, S. S. W., & Mervielde, I. (2010). Temperament, personality and developmental psychopathology: A review based on the conceptual dimensions underlying childhood traits. *Child Psychiatry and Human Development, 41*, 313–329.
Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73*, 1246–1256.
Gerris, J. R. M., Houtmans, M. J. M., Kwaaitaal-Roosen, E. M. G., Schipper, J. C., Vermulst, A. A., & Janssens, J. M. A. M. (1998). *Parents, adolescents and young adults in Dutch families: A longitudinal study*. Nijmegen, The Netherlands: Institute of Family Studies, University of Nijmegen.
Goldberg, L. R. (1992). The development of markers for the Big Five factor structure. *Psychological Assessment, 4*, 26–42.
Gough, H. G. (1987). *California psychological inventory administrator's guide*. Palo Alto, CA: Consulting Psychologists Press.
Hampson, S. E., & Goldberg, L. R. (2006). A first large cohort study of personality trait stability over the 40 years between elementary school and midlife. *Journal of Personality and Social Psychology, 91*, 763.

Jackson, J. J., Hill, P. L., Payne, B. R., Roberts, B. W., & Stine-Morrow, E. A. L. (2012). Can and old dog learn (and want to experience) new tricks? Cognitive training increases openness to experience in older adults. *Psychology and Aging, 27*, 286–292.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five inventory—Versions 4a and 54*. Berkeley, CA: Institute of Personality and Social Research, University of California.

John, O. P., & Srivastava, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (Vol. 2, pp. 102–138). New York: Guilford Press.

Klimstra, T. A., Hale, W. W., III, Raaijmakers, Q. A., Branje, S. J., & Meeus, W. H. (2009). Maturation of personality in adolescence. *Journal of Personality and Social Psychology, 96*, 898.

Lodi-Smith, J., Jackson, J., Bogg, T., Walton, K. E., Wood, D., Harms, P., et al. (2010). Mechanisms of health: Education and health-related behaviors partially mediate the relationship between conscientiousness and self-reported physical health. *Psychology and Health, 25*, 305–319.

Marsh, H. W., Nagengast, B., & Morin, A. J. S. (2013). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology, 49*, 1194–1218.

McCrae, R. R., & Costa, P. T., Jr. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (Vol. 2, pp. 139–153). New York: Guilford Press.

McCrae, R. R., & Costa, P. T., Jr. (2003). *Personality in adulthood: A five-factor theory perspective* (2nd ed.). New York: Guilford Press.

McCrae, R. R., Costa, P. T., Jr., Ostendorf, F., Angleitner, A., Hrebickova, M., Avia, M. D., et al. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology, 78*, 173–186.

McCrae, R. R., Costa, P. T., Jr., Terracciano, A., Parker, W. D., Mills, C. J., De Fruyt, F., et al. (2002). Personality trait development from age 12 to age 18: Longitudinal, cross-sectional and cross-cultural analyses. *Journal of Personality and Social Psychology, 83*, 1456.

McCrae, R. R., Terracciano, A., & 78 Members of the Personality Profiles of Cultures Project. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology*, *88*, 547–561.

Mervielde, I., & De Fruyt, F. (1999). Construction of the Hierarchical Personality Inventory for Children (HiPIC). In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe. Proceedings of the eight European conference on personality psychology* (pp. 107–127). Tilburg, The Netherlands: Tilburg University Press.

Musek, J. (2007). A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality, 41*, 1213–1233.

Ozer, D. J., & Gjerde, P. F. (1989). Patterns of personality consistency and change from childhood through adolescence. *Journal of Personality, 57*, 483–507.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*, 322–338.

Prinzie, P., & Deković, M. (2008). Continuity and change of childhood personality characteristics through the lens of teachers. *Personality and Individual Differences, 45*, 82–88.

Pullmann, H., Raudsepp, L., & Allik, J. (2006). Stability and change in adolescents' personality: A longitudinal study. *European Journal of Personality, 20*, 447–459.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*, 353–387.

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126*, 3.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*, 1–25.

Roberts, B. W., & Wood, D. (2006). Personality development in the context of the neo-socioanalytic model of personality. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp. 11–39). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Saucier, G. (1994). Mini-markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment, 63*, 506–516.

Soto, C. J., & John, O. P. (2014). Traits in transition: The structure of parent-reported personality traits from early childhood to early adulthood. *Journal of Personality*, *82*, 182–199.

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology, 100*, 330–348.

Spearman, C. (1904). "General intelligence" objectively determined and measured. *American Journal of Psychology, 15*, 201–293.

Tackett, J. L., Krueger, R. F., Iacono, W. G., & McGue, M. (2008). Personality in middle childhood: A hierarchical structure and longitudinal connections with personality in late adolescence. *Journal of Research in Personality, 42*, 1456–1462.

Tackett, J. L., Slobodskaya, H. R., Mar, R. A., Deal, J., Halverson, C. F., Jr., Baker, S. R., et al. (2012). The hierarchical structure of childhood personality in five countries: Continuity from early childhood to early adolescence. *Journal of Personality, 80*, 847–879.

Tang, T. Z., DeRubeis, R. J., Hollon, S. D., Amsterdam, J., Shelton, R., & Schalet, B. (2009). Personality change during depression treatment: A placebo-controlled trial. *Archives of General Psychiatry, 66*, 1322–1330.

Tross, S. A., Harper, J. P., Osher, L. W., & Kneidinger, L. M. (2000). Not just the usual case of characteristics: Using personality to predict college performance and retention. *Journal of College Student Development, 41*, 323–334.

Wortman, J., Lucas, R. E., & Donnellan, M. B. (2012). Stability and change in the Big Five personality domains: Evidence from a longitudinal study of Australians. *Psychology and Aging, 27*, 867.

Part II
Psychosocial Skills: Key Constructs

Chapter 5
The Need for Cognition: Key Concepts, Assessment, and Role in Educational Outcomes

Andrew T. Jebb, Rachel Saef, Scott Parrigon, and Sang Eun Woo

5.1 Definitions and Background

The first scholarly introduction of *need for cognition* (hereafter NFC) is attributed to Cohen, Stotland, and Wolfe (1955) who conceptualized the concept as "a need to structure relevant situations in meaningful, integrated ways" (p. 291). Their focus was on the resulting anxiety and tension that occurs when individuals are unable to sufficiently interpret ambiguous experiences. Developing this notion into its contemporary form, Cacioppo and colleagues (e.g., Cacioppo & Petty, 1982; Cacioppo, Petty, & Kao, 1984; Cacioppo, Petty, Kao, & Rodriguez, 1986; Cacioppo, Petty, & Morris, 1983) defined NFC as an individual's "tendency to engage in and enjoy thinking" (Cacioppo & Petty, 1982, p. 116). Whereas Cohen et al. (1955) saw NFC in terms of tension and resolution, Cacioppo and colleagues focused on how individuals differ in their inclinations for, and enjoyment of, effortful cognitive activity, such as reasoning, information acquisition, and complex problem-solving. As such, the contemporary understanding of NFC moved away from its original formulation as a *need* toward an individual's personal *desire* or *inclination* for higher-level cognition (Cacioppo et al., 1986). This shift reflected a more general trend away in psychology from conceptualizing behaviors as motivated by "needs" and "drives." Nonetheless, the term *need for cognition* remained constant as a recognition of the original pioneering research that captured individual differences in cognitive motivation (e.g., Cacioppo et al., 1986; Cohen et al., 1955).

Supporting Cacioppo et al.'s conceptualization, research has shown that NFC explains systematic differences in individuals' engagement of cognitively effortful behaviors. For example, both Smith, Kerr, Markus, and Stasson (2001) and Petty,

A.T. Jebb • R. Saef • S. Parrigon • S.E. Woo (✉)
Department of Psychological Sciences, Purdue University,
703 Third Street, West Lafayette, IN 47907, USA
e-mail: ajebb@purdue.edu; rsaef@purdue.edu; sparrigo@purdue.edu; sewoo@purdue.edu

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_5

Cacioppo, and Kasmer (1985) found that individuals higher in NFC were more strongly motivated to work on cognitive tasks within groups in contrast to other individuals who engaged in social loafing. The results of Ferguson, Chung, and Weigold (1985) showed that, in comparison to individuals with low NFC, high NFC individuals were more likely to seek information about a wide range of tasks, issues, and current events. These individuals also consulted more intellectually stimulating formats when seeking this information: e.g., newspapers as opposed to television. Thus, NFC is strongly related to an individual's effort to make sense of their world. However, the nature and extent to which this occurs is highly variable across individuals and has numerous implications for learning and educational outcomes (Cacioppo, Petty, Feinstein, & Jarvis, 1996).

5.2 Relationships with Other Similar Psychosocial Constructs

NFC is a relatively broad construct that spans the domains of personality and intelligence assessment. It is therefore unsurprising that NFC has been associated with a number of psychosocial constructs in the psychology literature, including cognitive innovativeness (Venkatraman & Price, 1990), need for affect (Maio & Esses, 2001), self-esteem (Osberg, 1987), *low* dogmatism (Cacioppo & Petty, 1982), and *low* need for closure (Webster & Kruglanski, 1994). In this chapter, we focus on three constructs in particular which bear a notable conceptual similarity to NFC: *typical intellectual engagement* (Ackerman, Kanfer, & Goff, 1995), *openness to ideas* (McCrae & Sutin, 2009), and *epistemic curiosity* (Litman, 2008). Their conceptual overlap has been corroborated by empirical studies that found strong associations among measures that were designed to capture these key constructs (see Table 5.1 for the definitions and sample items of these constructs).

5.2.1 Typical Intellectual Engagement

NFC has been shown to have a strong association with *typical intellectual engagement* (TIE), defined as "an individual's aversion or attraction to tasks that are intellectually taxing" (Ackerman et al., 1995, p. 276). TIE is concerned with how an individual performs *on average* and was originally formulated in the intelligence literature as a response to the disproportionate focus on *maximal* intellectual performance (Goff & Ackerman, 1992). More often than not, an individual's cognitive functioning or performance is, to some degree, suboptimal; the conditions leading to the best possible performance are rarely all present. Regardless, theory and research on intelligence had been primarily interested in assessing an individual's

Table 5.1 Constructs related to NFC, definitions, and sample items

Construct	Definition	Sample items
Need for cognition	"The tendency to engage in and enjoy thinking" (Cacioppo & Petty, 1982, p. 116)	"I am not satisfied unless I am thinking"; "I often succeed in solving difficult problems that I set out to solve" (Cacioppo Petty, 1982, p. 121)
Typical intellectual engagement	"An individual's aversion or attraction to tasks that are intellectually taxing" (Ackerman et al., 1995, p. 276)	"I prefer my life to be filled with puzzles I must solve" (Goff & Ackerman, 1992, p. 540)
Openness to ideas	"Openness to ideas is seen not only in an active pursuit of intellectual interests for their own sake, but also in open-mindedness and a willingness to consider new, perhaps unconventional ideas" (Costa & McCrae, 1992, p. 17)	"Has high degree of intellectual capacity"; "Concerned with philosophical problems" (Costa & McCrae, 1995, p. 31)
Epistemic curiosity	The "desire for knowledge that motivates individuals to learn new ideas, eliminate information-gaps, and solve intellectual problems" (Litman, 2008, p. 1586)	"Thinking over new ideas and concepts is fun"; "I would like to understand how complicated things like computers work" (Litman & Spielberger, 2003)

potential for maximum performance, without considering how an individual tends to perform on average. The latter, being more common, is more practically relevant, and the formulation of TIE was intended to balance these two approaches in the understanding of intelligence at large.

Woo, Harms, and Kuncel (2007) found that TIE correlated very highly with NFC ($r = 0.78$) and that measures of these two constructs showed highly similar patterns of correlations with other external variables, such as *autonomous regulation in learning* and the *Big Five* personality factors. The authors stated that the findings implied a large degree of overlap between the two constructs and that future research was required in order to more fully explicate their relationship and determine if they are different enough to justify two separate concepts. Subsequently, Mussel (2010) found similarly high correlations between NFC and TIE (rs = 0.75 and 0.62), again suggesting a strong degree of relatedness if not redundancy. Taken together, these findings are not surprising given their strikingly similar definitions (TIE, "an individual's aversion or attraction to tasks that are intellectually taxing" Ackerman et al., 1995, p. 276; NFC, "a tendency to engage in and enjoy thinking" Cacioppo & Petty, 1982, p. 116) as well as the overlapping item contents in their respective measures (e.g., NFC, "I really enjoy a task that involves coming up with new solutions to problems" Cacioppo & Petty, 1982; TIE, "I prefer my life to be filled with puzzles I must solve" Goff & Ackerman, 1992). Thus, given their definitional similarity, related items, and empirically found correlations among measures, we conclude that the constructs of NFC and TIE are nearly indistinguishable and treat them accordingly in this chapter.

5.2.2 Openness to Ideas

Openness to ideas (OI) is commonly known as a facet of openness to experience, one of the dimensions of the Big Five model of personality (McCrae & Costa, 1987). Conceptual and empirical research has often linked the broad personality dimension of openness to NFC. For example, McCrae called NFC and openness "virtually the same construct" (2000, p. 269). Echoing other theorists (e.g., Fleischhauer et al., 2010), Soubelet and Salthouse (2010) noted the similarity of NFC and openness but also stated that "Openness/Intellect can be considered to be a broader personality trait that encompasses more dimensions (e.g., affective, sensory, attitudes, preferences) than the Need for Cognition disposition" (p. 3). Our own view is that while there is a considerable overlap between openness to experience and NFC (Woo et al., 2007), they are still meaningfully different in openness to experience that captures a wider content space. For example, items measuring openness are not limited to assessing traits like intellectual curiosity and engagement but also measure a broader array of personality dispositions, such as openness to *actions* and *aesthetics* (Woo et al., 2014). However, because OI is an openness facet related specifically to intellectual pursuits, it is likely to have a stronger connection with NFC, both empirically and conceptually.

The relationship between OI and NFC is intuitively plausible given that an individual's tendency toward—and enjoyment of—thinking depends on a prior willingness to receive novel ideas. That is, insofar as thinking involves striving for new insights, one's enjoyment of thinking depends on whether or not one is open to these new ideas. Thus, openness to ideas may be a prerequisite for high NFC. (Further, compare the NFC item shown above to the following item for OI: "Value intellectual matters"; Costa & McCrae, 1995, p. 31.) Empirical work supports this theoretical link. Berzonsky and Sullivan (1992) found a correlation of $r = 0.78$ between OI and NFC, and Mussel's (2010) study also yielded significant correlations (rs = 0.68 and 0.87). In Fleischhauer et al. (2010), facets of the Big Five factors were examined, and it was discovered that NFC correlated most strongly with OI (0.67). Although the authors raised concerns about a potential lack of discriminant validity between these two constructs, their confirmatory factor analysis yielded a substantially better fit to the data when these constructs were separated as opposed to when they were combined into a single factor. According to the authors, these two constructs may be conceptually distinct from each other in that OI may be restricted to intellectual contexts that contain *novelty*, whereas NFC is "a general tendency to actively invest cognitive resources *independent of context* [emphasis added]" (Fleischhauer et al., 2010, p. 94). Nonetheless, it is important to question (and carefully evaluate) whether drawing a theoretical distinction between two constructs with such highly correlated measures would provide meaningful insights for research and practice.

5.2.3 *Epistemic Curiosity*

NFC is also related to *curiosity*, "the desire for new knowledge or experience" (Litman & Silvia, 2006). Throughout the literature, various scales and empirical work on curiosity have emphasized different but related conceptualizations of this construct: e.g., as a facilitator of personal growth (Kashdan & Fincham, 2002; Kashdan, Rose, & Fincham, 2004), as motivation toward new opportunities for knowledge or sensory information (Collins, Litman, & Spielberger, 2004; Litman & Spielberger, 2003), and as a more global trait that disposes an individual to find many diverse topics interesting (Peterson & Seligman, 2004). Recognizing the potential overlap between measures of curiosity and NFC, a number of studies have attempted to empirically examine this relationship. Olson, Camp, and Fuller (1984) compared NFC scores with a variety of curiosity measures, finding significant correlations with specific curiosity ($r=0.50$; defined below), academic curiosity ($r=0.68$), Spielberg's Trait curiosity ($r=0.67$), and state curiosity ($r=0.55$).

More specifically, NFC has been often linked to *epistemic curiosity* (EC), defined as the "desire for knowledge that motivates individuals to learn new ideas, eliminate information-gaps, and solve intellectual problems" (Litman, 2008, p. 1586). Historically, epistemic curiosity has been separated into two dimensions: *specific* and *diverse* curiosity (Berlyne, 1960). The difference between the two can easily be deduced from their titles: Specific curiosity regards a desire for particular pieces of information whereas diverse curiosity reflects a more general motivation for intellectual exploration. The correlation between these dimensions is strong (0.56; Litman & Spielberger, 2003), suggesting that they might be representing a common construct (i.e., epistemic curiosity). Mussel (2010) examined the discriminant validity among measures of epistemic curiosity and the three aforementioned constructs: NFC, TIE (Goff & Ackerman, 1992), and openness to ideas (Costa & McCrae, 1992). In two samples, correlations between measures of NFC and EC ranged from 0.59 to 0.74 (mean $r=0.64$). Interpreting the aggregate results (i.e., the intercorrelations of these constructs and a subsequent exploratory factor analysis), the author concluded that discriminant validity could not be established between the various curiosity constructs and the three constructs mentioned above. This conclusion was qualified by the statement that potential differences may indeed exist. Nonetheless, these results indicate a degree of redundancy, implying the need for further exploration and potential integration among these interrelated constructs.

5.2.4 *Further Thoughts on the Relationships of NFC with TIE, OI, and EC*

Based on a series of meta-analytic path analysis findings, Von Stumm, Hell, and Chamorro-Premuzic (2011) stated that "need for cognition, epistemic curiosity, and TIE are exemplary representatives of a group of investment trait constructs that

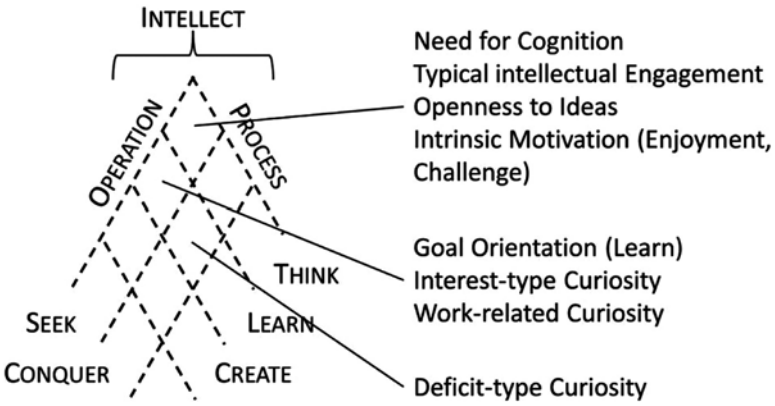

Fig. 5.1 Mussel's (2013) theoretical model depicting the relationship between intelligence and intelligence-based personality traits

describe tendencies to seek out, engage in, enjoy, and pursue opportunities for effortful cognitive activity" (p. 577). Delving further into their relatedness, a question that immediately arises is: "Are these constructs completely interchangeable, or are there any meaningful differences?" Mussel's (2013) study offers a number of important insights in regard to this question. Mussel (2013) proposed an integrative, two-dimensional structure of *Intellect*, defined as a dispositional variable that motivates intellectual behavior and mental processes (see Fig. 5.1). The two-dimensional structure comprises a process dimension and an operational dimension.

The process dimension contains two motivational orientations: *Seek*, the "affective aspects and a general openness that accompanies approaching situations that are intellectually challenging" and "the anticipated pleasure of new discoveries" (pp. 886–887), and *Conquer*, a motivation to be diligent, persistent, and "work hard to resolve incongruities and master intellectual challenges" once a situation is encountered (p. 886). The second dimension describes the psychological mechanisms that enable intellectual achievements. Mussel's model posits three types of operations: *think*, reflecting concepts of fluid intelligence and an appreciation of thinking about theories, problem-solving behaviors, and analyzing complex situations; *learn*, corresponding with concepts of crystallized intelligence, questioning/hypothesizing to fill knowledge gaps, a preference for gathering/understanding new knowledge, and an interest in knowing everything about a wide variety of topics; and *create*, "personality component contributing toward creative intellectual achievements"(p. 887). Crossing these two dimensions, Mussel proposed six distinct facets of intellect (i.e., seek-think, seek-learn, seek-create, conquer-think, conquer-learn, and conquer-create) and developed a new set of scales in order to capture these facets. Using multidimensional scaling, Mussel showed that scores of NFC, TIE, and OI fell within the *Seek-Think* facet of the structure. Scale scores reflecting EC, on the other hand, were positioned within the *Seek-Learn* facet (Mussel, 2013).

Based on these findings, we tentatively conclude that NFC (along with TIE and OI) mainly captures one's implicit proclivity for seeking out and enjoying the process of thinking itself, regardless of context, rather than for a more product-based motivation, such as that related to academic achievement and learning outcomes. In the following three sections, we build on this more nuanced understanding of NFC by presenting *empirical* and *practical* discussions around how NFC is assessed and relates to various learning and educational outcomes and how NFC may be cultivated in K-12 classrooms.

5.3 Measurement of NFC

The first published measure of NFC was Cacioppo and Petty's (1982) 34-item, self-report Need for Cognition Scale (NCS), which was subsequently developed into a shortened 18-item version (Cacioppo et al., 1984). The validity of both forms of the NCS has been supported by a number of studies that used both adult and adolescent samples (Cacioppo et al., 1996). The scales have also been translated into various languages (e.g., Bless, Wänke, Bohner, Fellhauer, & Schwarz, 1994). Furthermore, considering the relevance of NFC to various cognitive and academic outcomes (as discussed later), there is an important need to assess NFC in early stages of intellectual development (e.g., grade school); measuring NFC across these earlier life stages not only provides information about how NFC develops over time but may also help identify students that require additional academic attention. Although the original NFC scales were not designed to measure children, several studies have adapted the NFC self-report scales to younger developing populations as well (e.g., ages ten and older; Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Preckel, 2014).

5.3.1 The Dimensionality of NFC

Although the majority of research has yielded support for a single factor underlying NFC (Cacioppo et al., 1984, 1996; Cacioppo & Petty, 1982), several studies have indicated a *multidimensional* structure, raising the important question of whether or not NFC consists of subdimensions. For example, using a true/false format for NCS items, Tanaka, Panter, and Winborne (1988) identified three subscales within the original 34-item NFC scale: *cognitive persistence* (enjoyment of engaging in cognitive tasks), *cognitive confidence* (confidence about engaging in cognitive activities), and *cognitive complexity* (preference for complex or simple information-processing demands). Similarly, Davis, Severy, Kraus, and Whitaker (1993) identified two factors within the short-form scale: *enjoyment of cognitive effort* and *preference for problem-solving*. More recently, Furnham and Thorne (2013) also found evidence for multidimensionality of the 34-item scale with three distinct subdimensions: *motivation for cognitive challenge*, *motivation to gain knowledge and understanding*, and

enjoyment of thinking. As can be seen, the conceptual similarity of these factors found across independent studies is highly consistent. Therefore, while there is a significant amount of research supporting the idea that NFC is comprised of a single factor (Cacioppo et al., 1984, 1996; Cacioppo & Petty, 1982), there is also conflicting evidence for multidimensionality, a conflict that should be more fully investigated in future work.

5.3.2 NFC Measure Formats

The typical response format of NFC measures is self-report. In consideration of the limitations of self-report (e.g., biases and measurement artifacts related to wording), Fleischhauer, Strobel, Enge, and Strobel (2013) designed an *implicit* (i.e., behavioral, not self-report) measure and tested its validity against a self-report measure of NFC. Their results suggested that self-report measures better predict reflective, or controlled aspects of NFC-related behaviors, whereas an implicit measure added predictive validity for more spontaneous aspects of NFC-related behaviors. Conceptually, both aspects of NFC are important, as NFC is associated with both voluntary investment of cognitive effort and automatic attention allocated toward relevant stimuli *(*Enge, Fleischhauer, Brocke, & Strobel, 2008). In light of these findings, using both explicit (i.e., self-report) and implicit measures of NFC can provide a fuller assessment of the construct and can better predict automated behaviors motivated by NFC (e.g., Fazio, 1990; Strack & Deutsch, 2004*).* These findings foreshadow the impact of assessment method on the empirical relationships between NFC and various educational and intellectual outcomes (e.g., reading, learning, and academic performance) which may not be fully represented by simple self-report measures.

5.4 The Role of NFC (and Related Constructs) in Learning and Education

5.4.1 Reading

Research suggests that individuals with high NFC tend to be "good readers." Cacioppo et al. (1983) investigated how undergraduate students' levels of NFC moderated their ability to assess the logical consistency and strength of different arguments. Students with high NFC displayed an ability to discriminate between strong and weak arguments and were more influenced by the former. In contrast, students with low NFC did not significantly discriminate between the two and were not influenced by the quality of argumentation. Furthermore, students high in NFC were more apt to recall the argued information and reported longer periods of engaged thinking relative to those with low NFC.

Research regarding NFC's relationship to argument quality runs parallel to NFC's role in text comprehension. Dai and Wang (2007) tested university students'

comprehension of both narrative and expository texts after assessing individual differences in NFC. It was found that NFC was a significant predictor of comprehension across both genres, such that NFC accounted for 24 % and 34 % of text comprehension in narrative and expository texts, respectively. Interpreting these results, it may be that individuals high in NFC simply devote more cognitive energy when comprehending texts. Additionally, over time individuals high in NFC may be able to comprehend information more quickly and with greater facility than those low in this trait. Regardless, this study yielded support that NFC leads to a deeper engagement with presented information.

Similarly, Bråten, Anmarkrud, Brandmo, and Strømsø (2014) used a sample of 279 Norwegian secondary school students to show that NFC predicted text comprehension through the mediation of deeper-level strategies: Those high in NFC were more likely to engage in deeper-level strategies (i.e., trying to compare, contrast, and integrate across multiple texts), which in turn led to improved reading comprehension. This particular finding also points to the role of NFC in differentiating between deep versus surface learning approaches, as discussed in the following section.

5.4.2 Deep Versus Surface Learning (and Learning Motivation)

Deep and surface level learning represent two qualitatively different learning approaches that are well established in the literature (e.g., Ford, 1981; Newble & Entwistle, 1986). Surface learning refers to a more shallow type of information processing and integration (e.g., rote memorization). In contrast, deep learning involves a thorough conceptual understanding of the material and its relationship to other concepts. Theoretically, individuals high in NFC may be more disposed to deep learning, whereas NFC may be only minimally related to surface learning. In line with this hypothesis, Heijne-Penninga, Kuks, Hofman, and Cohen-Schotanus (2010) found that measures of deep learning were significantly related to NFC in a sample of medical students. Students who were high in NFC also performed better on both open- and closed-book tests. Similarly, previous research also suggests that openness is positively related to deep learning (Chamorro-Premuzic & Furnham, 2009; Duff, Boyle, Dunleavy, & Ferguson, 2004) and negatively related to surface learning (Chamorro-Premuzic, Furnham, & Lewis, 2007; Furnham, Christopher, Garwood, & Martin, 2007; Zhang, 2003).

TIE is also significantly related to deeper, more intrinsically motivated learning strategies and orientations (deep, mastery, and achievement orientations; Ackerman, Bowen, Beier, & Kanfer, 2001; Furnham, Monsen, & Ahmetoglu, 2009) and several learning interest measures (e.g., investigative interests, artistic interests; Ackerman et al., 2001). Furthermore, TIE was negatively related to "surface motivations," (i.e., attempting to barely meet the minimal requirements; Furnham et al., 2009), and positively related to autonomous self-regulation—the extent to which an individual is internally motived to learn (Woo et al., 2007). Similarly, Arteche, Chamorro-Premuzic, Ackerman, and Furnham (2009) posited that intrin-

sic motivation and TIE were precursors for deep learning and found that TIE was significantly and positively linked to deep learning.

5.4.3 Academic Performance

In practice, NFC has been modestly tied to academic performance among students. This aligns with studies that have shown that NFC is only marginally related to measures of general intelligence (e.g., $r = 0.15$, $r = 0.32$, and $r = –02$; see Cacioppo et al., 1996), as well as measures of reasoning in students in grades 2–9 (e.g., $r_{range} = –0.08$–0.25, Kokis et al., 2002; $r_{range} = –0.04$–20, Toplak, West, & Stanovich, 2014). However, NFC has been shown to be positively correlated with common metrics of academic performance, including high school grade point average ($r = 0.26$), undergraduate grade point average ($r = 0.14$), and the American College Test (ACT; $r = 0.20$; Petty & Jarvis, 1996). Other studies have reported similar associations between both the ACT (Cacioppo & Petty, 1984; Olson et al., 1984) and undergraduate GPA (Cacioppo & Petty, 1984; Petty & Jarvis, 1996; Tolentino, Curry & Leak, 1990; Waters & Zarkrajsek, 1990). Furthermore, Meier, Vogl, and Preckel (2014) found that NFC was the strongest predictor of attendance in classes for gifted students after other important individual differences were taken into account (e.g., academic self-concepts, academic interests, and mastery-and-performance goal orientations).

TIE has also displayed a strong relationship with various aspects of academic performance, such as course evaluations (i.e., grades; Chamorro-Premuzic, Furnham, & Ackerman, 2006a; Von Stumm, Hell, & Chamorro-Premuzi, 2011), math and English certification tests (Furnham et al., 2009), domain-specific knowledge and achievement (e.g., biology, humanities, civics, etc.; Ackerman et al., 2001; Chamorro-Premuzic, Furnham, & Ackerman, 2006b), and general knowledge (Ackerman et al., 2001; Chamorro-Premuzic et al., 2006a; Furnham et al., 2009; Ackerman & Heggestad, 1997). TIE has also shown consistent incremental validity above general measures of intelligence (i.e., IQ) and other personality measures in predicting academic outcomes (Chamorro-Premuzic et al., 2006b).

5.4.4 Interests and Attitudes

NFC has also been linked to how people develop and/or maintain interests in and attitudes toward various stimuli in the environment (e.g., information, objects, and other human beings). For example, the aforementioned study by Dai and Wang (2007) showed that NFC predicted higher topic interest through the effect of reading comprehension. In other words, readers with high NFC are more likely to develop an interest in a topic they read about because they tend to understand the materials better and at a deeper level.

A study by Kim, Yoon, Whang, Tversky, and Morrison (2007) suggested another interesting idea: that individuals high in NFC prefer less active stimuli because they have greater freedom in generating their own inferences about the object. The authors found that fourth-grade students with low NFC perceived *animated* presentations to be more comprehensible, interesting, and enjoyable, whereas those with high NFC tended to enjoy the presentation regardless of its type (i.e., animated or static). This may reflect a more general enjoyment of playing an active role in processing presented information.

Work conducted by Ackerman and colleagues on trait complexes (Ackerman, 1997; Ackerman & Heggestad, 1997; Ackerman & Beier, 2003; Ackerman et al., 2001) showed that openness and TIE were closely linked to artistic and investigative interests which were in turn associated with crystallized intelligence and ideational fluency. Using a sample of 655 introductory psychology undergraduate students, Feist (2012) found that NFC predicted interest in science ($r = 0.27$) even after controlling for the Big Five personality factors. O'Hara and colleagues (2009) collected data from 251 alumni and introductory psychology students in the Midwest and found that those with high NFC tended to show stronger political interest. Taking these findings altogether, the literature appears to suggest that NFC predicts individuals' interests in various content areas, especially in areas where deeper, more elaborate cognitive processing is required (e.g., sciences, advanced-level reading). As Schmidt (2014) stated, the construct *need for cognition* captures a "general interest in learning or knowledge acquisition" or a "broad interest in general learning—the proclivity to seek knowledge in a wide variety of knowledge areas" (p. 213).

NFC has also been shown to affect the persistence and resistance to attitude changes. For instance, Haugtvedt and Petty (1992) found that when individuals were exposed to a persuasive message about a novel, unfamiliar object and formed a certain attitude toward the object as a result (through an experimental manipulation), those with high NFC maintained the newly formed attitude more firmly over time compared to those with low NFC. In the same study, individuals were also exposed to negative information about a novel object and then given positive information that conflicted with their initially negative evaluation. High-NFC individuals displayed greater attitude resiliency given this conflicting message, indicating that altering their evaluations requires a greater amount of information. Given this finding, NFC has potentially important implications for understanding how students' attitudes and beliefs about a given subject or object may change or persist over time as a function of a relatively stable disposition.

5.5 Practical Implications: Cultivating NFC in Classroom

The present psychological literature is relatively silent on issues related to the development of NFC, such as how it may be nurtured through educational interventions, if there is a critical period for its development, and who can stand to benefit

the most in developing NFC (e.g., how NFC may compensate for deficiencies like low IQ). NFC has been traditionally studied as a stable, enduring dispositional characteristic and, as such, has been treated as an explanatory variable rather than as an outcome itself. This is not to say, however, that dispositional traits (e.g., intelligence, personality) cannot be developed over time. Indeed, personality researchers have long argued that personality evolves and matures over time, an idea that has been supported by numerous empirical studies (Roberts & DelVecchio, 2000; Roberts, Wood, & Caspi, 2008; Roberts & Mroczek, 2008). In general, personality has been found to be less stable during adolescence than it is later in life (Roberts & DelVecchio, 2000), suggesting increased malleability during this age period. Support for this same potential malleability in NFC has been found by Padgett et al. (2010) who showed that NFC increased in first-year undergraduates who had interactions with faculty outside of the classroom and who held meaningful discussions with diverse peers, suggesting that an intellectual stimulating environment might be able to increase NFC. Additionally, research has found that variation in mean levels of the specific items of NFC exists across grades five and six, suggesting that there also may exist some level of change in the mean levels of NFC during this developmental period and that it might be similarly malleable (Preckel, 2014). Taken together, these results suggest that, while cultivating NFC potentially challenging, it is certainly not impossible.

How then can NFC be cultivated, encouraged, and trained in the classroom? How can students learn to be more intellectually curious, open to new ideas, and interested in engaging in effortful cognitive activities, such as solving complex math problems, learning scientific principles, and reading, understanding, and analyzing advanced texts? While one potential avenue was noted above (i.e., intellectual discussions with faculty and peers; Padgett et al., 2010), there may be a number of different methods that can lead to these desired ends. One potential strategy is based on the literatures of self-efficacy (Bandura, 1997; Elias & Loomis, 2002; see also Chaps. 8 and 10) and goal orientation (e.g., approach vs. avoid). Academic self-efficacy refers to a student's perception of their ability to successfully complete a task or attain a goal. Students with high self-efficacy are more engaged and work harder and longer when faced with a difficult task—behaviors that are strikingly similar to those related to high NFC (Bandura, 1997). Unfortunately, low self-efficacy may encourage students to be more passive during these tasks and thus forgo the opportunity to recognize and develop their abilities. If one does not feel capable of completing a task or gleaning intended information (low self-efficacy), they may elect to become more passive in their contribution or avoid the task all together, potentially impairing their academic performance. Therefore, high self-efficacy may be a perquisite for the pursuit and enjoyment of effortful thought (i.e., NFC), as it is difficult to see how someone would seek out and enjoy thinking if they do not believe themselves to be capable of successfully completing or attaining the target goal. Increasing a student's sense of self-efficacy could, therefore, encourage them to approach cognitively engaging and difficult tasks and, when successful in their pursuits, contribute to a stronger sense of academic self-efficacy. This has the ultimate effect of encouraging future engagements with difficult and complex

goals and tasks. Therefore, while these efficacy-related constructs are meaningfully distinguished from NFC-related constructs (Elias & Loomis, 2002; Von Stumm et al., 2011), they are certainly related. If a child has confidence in their ability to accomplish a task or academic goal, the student may select more difficult and complex goals, tasks, and classes and be more willing to engage in some of the cognitively demanding activities cited above, such as complex problem-solving. In line with this theorizing, Heppner, Reeder, and Larson (1983) found that undergraduate students high in NFC gave more positive appraisals of their problem-solving skills than students with low NFC. Positive beliefs regarding one's efficacy or competence have also been found to be related to students' reading and writing abilities and skills (Shell et al., 1989).

NFC may also be cultivated by programs aimed at increasing both the depth and frequency of students' reading. Reading ability has long been related to both increased academic success and various metrics of intelligence (McCullough, 1939; Wheeler & Wheeler, 1949), and reading frequency has significant consequences for the development of future cognitive skills (Cunningham & Stanovich, 1998). This is because an increase in reading ability enables word recognition to become more automated and therefore less cognitively demanding, leading to greater proficiency with complex syntactic structures, a more robust vocabulary, and an increase in general verbal skills (Cunningham & Stanovich, 1998). As research has shown a substantial relationship between perceived competence in an activity and one's enjoyment of it (e.g., Harackiewicz, Sansone, & Manderlink, 1985; Reeve, 1989), the increased cognitive benefits that result from reading are likely to lead individuals to enjoy other cognitively effortful activities—whether they are demanding in similar or distinct ways. Thus, implementing educational programs that promote reading both in and outside the classroom may lead to numerous cognitive and intellectual benefits, hailing from the development of not only NFC but many other important traits that can be cultivated through such practices.

5.6 Conclusion

The present literature on NFC offers voluminous information on how it is correlated to other theoretically similar constructs, such as TIE, curiosity, and openness to ideas. Although they are all concerned with intellectual engagement and the pursuit of information, these constructs seem to focus on and emphasize different aspects of these processes. Examining correlations among these related constructs is informative, but what is most needed at present is further *theoretical development* that delineates the interrelationships among these constructs and how their empirical redundancy can be reduced through possible integration. Establishing an overarching framework of these traits is essential for understanding not only the differences in intellectually related behaviors seen across individuals but for how individuals engage and interact with information more generally.

Further, research on developing or *cultivating* NFC is also lacking. Although changes in personality typically occur over longer periods of time, it is important to explore if there are any ways to promote this trait, as it represents an individual's broad relationship with challenging intellectual material that is related to academic outcomes and may help identify students who require additional educational attention. Most of the scales measuring NFC have been adapted to older adolescents (e.g., college students) or adults. However, as stated, NFC may be most relevant to younger children and further work needs to clarify the role it plays at this critical developmental stage.

NFC is important for describing, explaining, and predicting the intellectual pursuits and outcomes of individuals throughout different life stages. That is, individuals high in NFC are more likely to excel in academic environments, experience a greater ability to acquire information, think more critically, are better at assessing the quality of arguments, and hold attitudes that are less susceptible to fluctuations. Given the relevance of NFC for these critical outcomes, it is unfortunate that the vast majority of extant research has used measures of NFC to explain other educational outcomes (e.g., GPA, reading ability), as opposed to investigating NFC as an outcome *in and of itself*. Several potential strategies of how NFC may be cultivated were delineated above—viz., increasing students' reading frequency and sense of self-efficacy. The present task now becomes to test these theoretical foundations in order to determine their validity and any potential intervening (i.e., moderating) factors. Given the known importance of NFC, these questions are paramount and can substantially illuminate how NFC and other traits that dispose students toward intellectual pursuits and can be cultivated both within and outside the classroom.

References

Ackerman, P. L. (1997). Personality, self-concept, interests, and intelligence: which construct doesn't fit? *Journal of Personality, 65*, 171–204.

Ackerman, P. L., & Beier, M. E. (2003). Intelligence, personality, and interests in the career choice process. *Journal of Career Assessment, 11*(2), 205–218.

Ackerman, P. L., Bowen, K. R., Beier, M., & Kanfer, R. (2001). Determinants of individual differences and gender differences in knowledge. *Journal of Educational Psychology, 93*(4), 797–825.

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality and interests: Evidence for overlapping traits. *Psychological Bulletin, 121*, 219–245.

Ackerman, P. L., Kanfer, R., & Goff, M. (1995). Cognitive and noncognitive determinants and consequences of complex skill acquisition. *Journal of Experimental Psychology: Applied, 1*, 270–304.

Arteche, A., Chamorro-Premuzic, T., Ackerman, P., & Furnham, A. (2009). Typical intellectual engagement as a byproduct of openness, learning approaches, and self-assessed intelligence. *Educational Psychology, 29*, 357–367.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman and Company.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity* (McGraw-Hill series in psychology). New York: McGraw-Hill Book Company.

Berzonsky, M. D., & Sullivan, C. (1992). Social-cognitive aspects of identity style: Need for cognition, experiential openness, and introspection. *Journal of Adolescent Research, 7*, 140–155.

Bless, H., Wänke, M., Bohner, G., Fellhauer, R. F., & Schwarz, N. (1994). Need for cognition—A scale measuring engagement and happiness in cognitive tasks. *Zeitschrift für Sozialpsychologie, 25*, 147–154.

Bråten, I., Anmarkrud, Ø., Brandmo, C., & Strømsø, H. I. (2014). Developing and testing a model of direct and indirect relationships between individual differences, processing, and multiple-text comprehension. *Learning & Instruction, 30*, 9–24.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*, 116–131.

Cacioppo, J. T., & Petty, R. E. (1984). The need for cognition: Relationship to attitudinal processes. In R. P. McGlynn, J. E. Maddux, C. D. Stoltenberg, & J. H. Harvey (Eds.), *Social perception in clinical and counseling psychology* (pp. 91–119). Lubbock, TX: Texas Tech University.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*(2), 197–253.

Cacioppo, J. T., Petty, R. E., & Kao, C. E. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*, 306–307.

Cacioppo, J. T., Petty, R. E., Kao, C. F., & Rodriguez, R. (1986). Central and peripheral routes to persuasion: An individual difference perspective. *Journal of Personality and Social Psychology, 51*(5), 1032–1043.

Cacioppo, J. T., Petty, R. E., & Morris, K. J. (1983). Effects of need for cognition on message evaluation, recall, and persuasion. *Journal of Personality and Social Psychology, 45*(4), 805–818.

Chamorro-Premuzic, T., & Furnham, A. (2009). Mainly openness: The relationship between the Big Five personality traits and learning approaches. *Learning and Individual Differences, 19*, 524–529.

Chamorro-Premuzic, T., Furnham, A., & Ackerman, P. (2006a). Ability and personality correlates of general knowledge. *Personality and Individual Differences, 41*(3), 419–429.

Chamorro-Premuzic, T., Furnham, A., & Ackerman, P. (2006b). Incremental validity of the typical intellectual engagement scale as predictor of different academic performance measures. *Journal of Personality Assessment, 87*(3), 261–268.

Chamorro-Premuzic, T., Furnham, A., & Lewis, M. (2007). Personality and approaches to learning predict preference for different teaching methods. *Learning and Individual Differences, 17*(3), 241–250.

Cohen, A. R., Stotland, E., & Wolfe, D. M. (1955). An experimental investigation of need for cognition. *Journal of Abnormal and Personality Psychology, 51*, 291–294.

Collins, R., Litman, J., & Spielberger, C. (2004). The measurement of perceptual curiosity. *Personality and Individual Differences, 36*(5), 1127–1141.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the revised NEO personality inventory. *Journal of Personality Assessment, 64*, 21–50.

Cunningham, A. E., & Stanovich, K. E. (1998). What reading does for the mind. *American Educator, 22*, 8–15.

Dai, D., & Wang, X. (2007). The role of need for cognition and reader beliefs in text comprehension and interest development. *Contemporary Educational Psychology, 32*(3), 332–347.

Davis, T. L., Severy, L. J., Kraus, S. J., & Whitaker, J. M. (1993). Predictors of sentencing decisions: The beliefs, personality variables, and demographic factors of juvenile justice personnel. *Journal of Applied Social Psychology, 23*, 451–477.

Duff, A., Boyle, E., Dunleavy, K., & Ferguson, J. (2004). The relationship between personality, approach to learning and academic performance. *Personality and Individual Differences, 36*, 1907–1920.

Elias, S. M., & Loomis, R. J. (2002). Utilizing need for cognition and perceived self-efficacy to predict academic performance. *Journal of Applied Social Psychology, 32*, 1687–1702.

Enge, S., Fleischhauer, M., Brocke, B., & Strobel, A. (2008). Neurophysiological measures of involuntary and voluntary attention allocation and dispositional differences in need for cognition. *Personality and Social Psychology Bulletin, 34*, 862–874.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior—The mode model as an integrative framework. *Advances in Experimental Social Psychology, 23*, 75–109.

Feist, G. J. (2012). Predicting interest in and attitudes toward science from personality and need for cognition. *Personality and Individual Differences, 52*, 771–775.

Ferguson, M., Chung, M., & Weigold, M. (1985). *Need for cognition and the medium dependency components of reliance and exposure.* Paper presented at the meeting of the International Communication Association, Honolulu, HI.

Fleischhauer, M., Enge, S., Brocke, B., Ullrich, J., Strobel, A., & Strobel, A. (2010). Same or different? Clarifying the relationship of need for cognition to personality and intelligence. *Personality and Social Psychology Bulletin, 36*, 82–96.

Fleischhauer, M., Strobel, A., Enge, S., & Strobel, A. (2013). Assessing implicit cognitive motivation: Developing and testing an implicit association test to measure need for cognition. *European Journal of Personality, 27*, 15–29.

Ford, N. (1981). Recent approaches to the study and teaching of 'effective learning' in higher education. *Review of Educational Research, 51*, 345–377.

Furnham, A., Christopher, A. N., Garwood, J., & Martin, G. N. (2007). Approaches to learning and the acquisition of general knowledge. *Personality and Individual Differences, 43*, 1563–1571.

Furnham, A., Monsen, J., & Ahmetoglu, G. (2009). Typical intellectual engagement, Big Five personality traits, approaches to learning and cognitive ability predictors of academic performance. *British Journal of Educational Psychology, 79*, 769–782.

Furnham, A., & Thorne, J. (2013). Need for cognition: Its dimensionality and personality and intelligence correlates. *Journal of Individual Differences, 34*, 230–240.

Goff, M., & Ackerman, P. L. (1992). Personality-intelligence relations: Assessing typical intellectual engagement. *Journal of Educational Psychology, 84*, 537–552.

Harackiewicz, J. M., Sansone, C., & Manderlink, G. (1985). Competence, achievement orientation, and intrinsic motivation: A process analysis. *Journal of Personality and Social Psychology, 48*, 493–508.

Haugtvedt, C. P., & Petty, R. E. (1992). Personality and persuasion: Need for cognition moderates the persistence and resistance of attitude changes. *Journal of Personality and Social Psychology, 63*, 308–319.

Heijne-Penninga, M., Kuks, J. B. M., Hofman, W. H. A., & Cohen-Schotanus, J. (2010). The influence of need for cognition, deep learning and preparation time on open and closed-book test scores. *Medical Education, 44*, 884–891.

Heppner, P. P., Reeder, B. L., & Larson, L. M. (1983). Cognitive variables associated with personal problem solving appraisal: Implications for counseling. *Journal of Counseling Psychology, 30*, 537–545.

Kashdan, T.B. & Fincham, F.D. (2002) Facilitating creativity by regulating curiosity. *American Psychologist, 57*, 373–374.

Kashdan, T. B., Rose, P., & Fincham, F. D. (2004). Curiosity and exploration: Facilitating positive subjective experiences and personal growth opportunities. *Journal of Personality Assessment, 82*, 291–305.

Kim, S., Yoon, M., Whang, S. M., Tversky, B., & Morrison, J. B. (2007). The effect of animation on comprehension and interest. *Journal of Computer Assisted Learning, 23*(3), 260–270.

Kokis, J. V., Macpherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology, 83*, 26–52.

Litman, J. A. (2008). Interest and deprivation factors of epistemic curiosity. *Personality and Individual Differences, 44*, 1585–1595.

Litman, J. A., & Silvia, P. J. (2006). The latent structure of trait curiosity: Evidence for interest and deprivation curiosity dimensions. *Journal of Personality Assessment, 86*, 318–328.

Litman, J. A., & Spielberger, C. D. (2003). Measuring epistemic curiosity and its diversive and specific components. *Journal of Personality Assessment, 80*, 75–86.

Maio, G. R., & Esses, V. M. (2001). The need for affect: Individual differences in the motivation to approach or avoid emotions. *Journal of Personality, 69*, 583–615.

McCrae, R. R. (2000). Emotional intelligence from the perspective of the five-factor model of personality. In R. Bar-On & J. D. A. Parker (Eds.), *Handbook of emotional intelligence* (pp. 263–276). San Francisco, CA: Jossey-Bass.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81–90.

McCrae, R. R., & Sutin, A. R. (2009). Openness to experience. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 257–273). New York: Guilford.

McCullough, C. M. (1939). Relationship between intelligence and gains in reading ability. *Journal of Educational Psychology, 30*(9), 688–692.

Meier, E., Vogl, K., & Preckel, F. (2014). Motivational characteristics of students in gifted class: The pivotal role of need for cognition. *Learning and Individual Differences, 33*, 39–46.

Mussel, P. (2010). Epistemic curiosity and related constructs: Lacking evidence of discriminant validity. *Personality and Individual Differences, 49*, 506–510.

Mussel, P. (2013). Intellect: A theoretical framework for personality traits related to intellectual achievements. *Journal of Personality & Social Psychology, 104*(5), 885–906.

Newble, D. I., & Entwistle, N. J. (1986). Learning styles and approaches: Implications for medical education. *Medical Education, 20*, 162–175.

O'Hare, R. E., Walter, M. I., & Christopher, A. N. (2009). Need for cognition and conscientiousness as predictors of political interest and voting strategy. *Journal of Applied Social Psychology, 39*, 1397–1416.

Olson, K. R., Camp, C. J., & Fuller, D. (1984). Curiosity and need for cognition. *Psychological Reports, 54*, 71–74.

Osberg, T. M. (1987). The convergent and discriminant validity of the need for cognition scale. *Journal of Personality Assessment, 51*, 441–450.

Padgett, R. D., Goodman, K. M., Johnson, M. P., Saichaie, K., Umbach, P. D., & Pascarella, E. T. (2010). The impact of college student socialization, social class, and race on need for cognition. *New Directions For Institutional Research, 145*, 99–111.

Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. Washington, DC: American Psychological Association.

Petty, R. E., Cacioppo, J. T., & Kasmer, J. A. (1985). *Individual differences in social loafing on cognitive tasks*. Paper presented at the meeting of the Midwestern Psychological Association, Chicago.

Petty, R. E., & Jarvis, W. B. G. (1996). An individual differences perspective on assessing cognitive processes. In N. Schwarz & S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research* (pp. 221–257). San Francisco, CA: Jossey-Bass.

Preckel, F. (2014). Assessing need for cognition in early adolescence: Validation of a German adaption of the Cacioppo/Petty scale. *European Journal of Psychological Assessment, 30*(1), 65–72.

Reeve, J. (1989). The interest-enjoyment distinction in intrinsic motivation. *Motivation and Emotion, 13*, 83–103.

Roberts, B., & DelVecchio, W. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126*, 3–25.

Roberts, B., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science, 17*(1), 31–35.

Roberts, B. W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 375–398). New York: Guilford.

Schmidt, F. (2014). A general theoretical integrative model of individual differences in interests, abilities, personality traits, and academic and occupational achievement: A commentary on four recent articles: Integrative comments on abilities, interests, and achievements. *Perspectives on Psychological Science, 9*, 211–218.

Shell, D. F., Murphy, C. C., & Bruning, R. H. (1989). Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. *Journal of Educational Psychology, 81*, 91–100.

Smith, B. N., Kerr, N. A., Markus, M. J., & Stasson, M. F. (2001). Individual differences in social loafing: Need for cognition as a motivator in collective performance. *Group Dynamics, Theory, Research, and Practice, 5*, 150–158.

Soubelet, A., & Salthouse, T. (2010). The role of activity engagement in the relations between openness/intellect and cognition. *Personality and Individual Differences, 49*, 896–901.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*, 220–247.

Tanaka, J. S., Panter, A. T., & Winborne, W. C. (1988). Dimensions of the need for cognition: Subscales and gender differences. *Multivariate Behavioral Research, 23*, 35–50.

Tolentino, E., Curry, L., & Leak, G. (1990). Further validation of the short form of the need for cognition scale. *Psychological Reports, 66*, 321–322.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Rational thinking and cognitive sophistication: Development, cognitive abilities, and thinking dispositions. *Developmental Psychology, 50*(4), 1037–1048. doi:10.1037/a0034910.

Venkatraman, M. P., & Price, L. L. (1990). Differentiating between cognitive and sensory innovativeness: Concepts, measurement, and implications. *Journal of Business Research, 20*, 293–315.

Von Stumm, S., Hell, B., & Chamorro-Premuzi, T. (2011). The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science, 6*, 574–588.

Waters, L. K., & Zarkrajsek, T. (1990). Correlates of need for cognition total and subscale scores. *Education and Psychological Measurement, 50*(2), 13–17.

Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology, 67*, 1049–1062.

Wheeler, L. R., & Wheeler, V. D. (1949). The relationship between reading ability and intelligence among university freshmen. *Journal of Educational Psychology, 40*(4), 230–238. doi:10.1037/h0062904.

Woo, S. E., Chernyshenko, O., Longley, A., Zhang, Z., Chiu, C., & Stark, S. (2014). Openness to experience: Its lower level structure, measurement, and cross-cultural equivalence. *Journal of Personality Assessment, 96*, 29–45.

Woo, S. E., Harms, P. D., & Kuncel, N. R. (2007). Integrating personality and intelligence: Typical intellectual engagement and need for cognition. *Personality and Individual Differences, 43*, 1635–1639.

Zhang, L.-f. (2003). Does the big five predict learning approaches? *Personality and Individual Differences, 34*, 1431–1446.

Chapter 6
Understanding Creativity in the Schools

James C. Kaufman, Ronald A. Beghetto, and Anna Dilley

6.1 Understanding Creativity in the Schools

Creativity in the schools is many things to many people – a promise, a threat, a hope, a distraction, or a goal. It is seen as important, yet often ignored or back-burnered (Aljughaiman & Mowrer-Reynolds, 2005; Beghetto & Plucker, 2006). General interest in making creativity a more central goal in schools has grown steadily over the years, as it has been covered by TED talks, layperson magazines, popular blogs, and academic books. However, unless educators understand what creativity is and how to foster it, it is unlikely that interest in creativity will develop into approaches that actually foster students' creative potential (Beghetto & Kaufman, 2013; Kaufman & Beghetto, 2014).

The main purpose of this chapter is to provide an overview of what creativity means in a classroom context. This includes providing insights into how it can be fostered and assessed. We open the chapter by discussing how creativity is typically defined. We introduce the Four-C Model of Creativity, which we believe is helpful for understanding creativity in and out of school contexts. We then discuss the various ways creativity can and has been assessed. Next we discuss various insights for how educators might foster students' creative potential. We close by discussing how creative metacognition may help students express their creativity in the classroom.

J.C. Kaufman (✉) • R.A. Beghetto • A. Dilley
Educational Psychology, Neag School of Education, University of Connecticut,
2131 Hillside Road, Storrs, CT 06269-3007, USA
e-mail: James.kaufman@uconn.edu; Ronald.beghetto@uconn.edu;
Anna.Dilley@uconn.edu

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_6

6.2 Creativity: Definitions and Core Concepts

Creativity research did not blossom until the 1950s, after Guilford gave a now legendary presidential address at the American Psychological Association (Guilford, 1950). He argued for creativity's importance as both a scholarly pursuit and real-world outcome (Kaufman, 2009). Guilford identified schools as a key context for cultivating creativity – stressing the importance of nurturing the "creative promise" of our children. Since that time, creativity has developed as a field of study across numerous disciplines (such as psychology, education, and business) and around the world.

Recognizing the importance of creativity is one thing; understanding what creativity means in a classroom context is quite another (Kaufman & Beghetto, 2013b). Creativity researchers have long agreed that creativity comprises two key elements (see, e.g., Barron, 1955; Guilford, 1950; Stein, 1953). The first element is that creativity must be something new or different. But originality is not enough; for something to be creative, it must also be task appropriate or useful. Simonton (2012) frames it as Creativity = Originality × Appropriateness ($C = [O \times A]$). If something is not appropriate or not original, then it cannot be considered creative. We have briefly elaborated on this expression to include context: $C = [O \times A]_{context}$ (Beghetto & Kaufman, 2014). This elaboration highlights the fact that originality and appropriateness are determined by a particular social, cultural, and historical context. One of the theories we will discuss in the pages to come, Csikszentmihalyi's (1999) systems model, is established in the interplay between the creator and the larger context of the gatekeepers and the domain.

In addition to this definition, researchers have also outlined useful ways of categorizing creativity. Rhodes (1962), for instance, proposed four Ps of creativity – person, process, product, and press (e.g., environment). The *person* is the creator; research on this topic can include which personal factors influence creativity (e.g., openness to experience, sensible risk taking, and confidence in one's ideas). The creative *process* is the actual process of being creative (e.g., generating, implementing, and evaluating ideas). The creative *product* refers to the outputs that may be considered (e.g., ideas, behaviors, material objects); research on this topic may also cover the ways that people evaluate creative work. Finally, research on the *press* can focus on which factors of the environment can support or stifle creativity. The four Ps of creativity can thereby be helpful for educators in thinking about the various interrelated elements necessary for nurturing students' creative potential.

Let's explore an example. Let's say that a teacher (Mr. Herman) wants to help a student (Nicole) be more creative. He might want to first think of which of the four Ps he wants to emphasize. Does Mr. Herman want to help Nicole's "person" element and teach her strategies to help her take appropriate risks and appreciate a large array of potential ideas and possibilities? Does he want to instead think about how to improve Nicole's creative "products" by working with her on specific projects and giving her constructive feedback? Or, perhaps, he may choose to focus on the "process" and teach Nicole strategies that will help her be creative (such as improving her divergent thinking; see Beghetto, Kaufman, & Baer, 2015).

Conversely, Mr. Herman may take a step back and look at the surrounding "press" – he may ask himself if he is leading a classroom that allows students to express their creativity without worrying about their ideas being dismissed or insulted.

In more recent years, researchers have elaborated on these four Ps. Simonton (1990), for instance, added *persuasion* – highlighting the importance of successful creators being able to articulate and persuade others of the merit of their creative expression. Runco (2003) added *potential*, which is particularly germane to educational settings as it stresses the importance of recognizing yet-to-be fulfilled creative potential. Glaveanu (2013) further elaborated on the various Ps – emphasizing how social and cultural context plays a role in creativity. Glaveanu has conceptualized the Ps into five As. The person, for instance, becomes the *actor*, the process becomes the *action*, and the product becomes the *artifact*. Press is split into two concepts: *audience* and *affordances*. The audience refers to the people who respond to the creative work (from immediate peers to a global network of millions), and affordances are the specific materials (or lack thereof) that a person can access. These elaborations are helpful in that they more adequately capture the complexity of creativity. Be they Ps or As, these categories can serve as useful reminders that creativity is a multifaceted phenomenon that requires simultaneous consideration of individuals and contexts.

There are two key theories that primarily focus on the creative person. They offer a synthesis of important individual differences that influence a person's creativity. One is the investment theory of creativity (Sternberg & Lubart, 1995). This theory is based on the idea that creative people "buy" bad ideas at a low price and then "sell" them to others after they have improved them. The creative person will then generally move on to the next "bad" idea. This theory of creativity illustrates the point that for an idea to be considered creative, it cannot simply be more original than the next idea; the creator also needs to sell it to their audience.

According to Sternberg and Lubart, there are six factors that comprise the basis for creative thought. Three of these factors are constructs that are studied even more than creativity – intelligence, personality, and motivation. Intelligence and creativity have been found to be related (Kim, 2005), although at a relatively low level. We've briefly discussed personality (creative people are most notably open to experience; Kaufman, 2009); it is also worth noting that creative people are not necessarily more likely to be mentally ill – this connection is overstated and complex (see essays in Kaufman, 2014). We will address motivation later in the chapter. The fourth factor, environment, is one of the four Ps and the only one of the six that is not person driven. The fifth factor is thinking styles, which are the way that people prefer to use their intellectual abilities (see essays in Zhang, Sternberg, & Rayner, 2012). A person's thinking style may, for example, tend toward analysis, practicality, or creativity (Sternberg & Zhang, 2001). The final factor, knowledge, refers to specific information in a domain. A recurring debate in creativity is the importance of the domain – whether someone who is creative in poetry, for example, is also more likely to also be creative in mathematics or business (Baer & Kaufman, 2006; Kaufman & Baer, 2004). One key nuance to this debate is the importance of knowledge to creativity; someone who is a creative poet has a vast amount of

knowledge about writing poetry (such as an understanding of rhythm, meter, and rhyme) that cannot be easily applied to other domains. That is, creativity seems to be much more domain specific than, for example, intelligence.

If someone has a deficit in one or more of these six factors, then that person would likely be less creative. For example, if a person is not motivated to try out new ideas, they might not feel any need to think creatively.

Another theory grounded in the idea that creativity has many different facets is the Componential Model of Creativity (Amabile, 1996). Amabile states that for creativity to occur, an individual must have three components: domain-relevant skills (including knowledge, expertise, technical skills, intelligence, and more), creativity-relevant processes (such as personality, cognitive style, propensity for risk taking, etc.), and task motivation (an intrinsic desire to solve the problem or finish the task). For example, for painters to be optimally creative, they must have domain-relevant skills (such as knowledge of color and brushwork), creativity-relevant processes (such as a personality that is willing or even eager to face scrutiny for going against the grain), and task motivation (e.g., the desire to express themselves emotionally after a trauma or simple love of the process of painting).

Another way to approach the question of creativity that is particularly related to K-12 education is the Four-C Model (Beghetto & Kaufman, 2007; Kaufman & Beghetto, 2009), which views creativity as a developmental progression from everyday creativity to creative eminence. It branches off from the core distinction between "little-c" (everyday creativity) and "Big-C" (eminent creativity). We will highlight approaches that center around one of these two conceptions, and then we will outline the extra dimensions of the Four-C Model (i.e., Pro-c and mini-c creativity).

6.3 The Four-C Model

6.3.1 Big-C Approaches

When most people think of creativity, they tend to take one of two perspectives. The first focuses on legendary creativity – genius work that has stood the test of time. These types of studies and theories are usually referred to as studying "Big-C" creativity. One of the premiere Big-C researchers is Simonton (2009). Some examples of his work include studies that examine the creative contributions of classical and opera composers whose works have lasted centuries (e.g., Simonton, 1977, 1997), or renowned scientists (Simonton, 2004). Much Big-C research uses the historiometric method, which analyzes data derived from biographies or other reference sources, thus allowing researchers to study cross-group patterns in a manner that would be impossible to do if they were interviewing each person individually. Indeed, most creative geniuses are either deceased or not interested in answering surveys!

Various creativity theories have focused on Big-C conceptions of creativity. Csikszentmihalyi's (1999) systems model, for instance, describes eminent creativity as resulting at the interaction between domain, field (i.e., gatekeepers), and person. This theory is helpful for explaining how creative contributions attain legendary status. Big-C status is typically bestowed on creators or works at some later point in time when critics, connoisseurs, and domain experts recognize the creative talent inherent in a body of work.

The propulsion theory of creative contributions (Sternberg & Kaufman, 2012; Sternberg, Kaufman, & Pretz, 2001, 2002, 2003; Sternberg, Pretz, & Kaufman, 2003) is another example. This theory focuses on how an individual's creative act potentially changes an entire field. There are eight distinct types of creative contributions included in the theory. The first four types stay within the core paradigms in the field, whereas the next four are more radical shifts. First is *replication*, in which the creator simply reproduces past work in their field; examples in the movie industry include the reboots of classic movies, such as the remakes of *Footloose*, *Night of the Living Dead*, *Flight of the Phoenix*, and *Total Recall*. Second is *redefinition*, which differs from replication in that it adds a new twist to a classic idea. An example might be Jane Austen novels told in the form of a video blog. Next is *forward incrementation*, in which the creator moves his or her field of study forward, but not drastically. Many scientific research papers fall into this category as replication can be seen as a cornerstone of science – classic studies are reworked with one key change (replication-extension studies; Bonett, 2012; i.e., the participants are children instead of adults). The final category that stays within the paradigm is *advance forward incrementation*. Here the field is moved forward drastically, often leaving the creator in the difficult position of not attracting an audience given that they are "ahead of their time." Many movies that have attained a "cult" status but were not initial hits may represent advance forward incrementation (such as *Scott Pilgrim vs the World* or *Rocky Horror Picture Show*).

In contrast, the next four categories of creative contributions are the attempts to truly change the face of their chosen field. The first, *redirection,* is the attempt of a creator to move the field in a completely different direction, such as Galileo proposing his theory of heliocentricity. *Reconstruction/redirection* is an attempt to move a field back to its starting point, in order to start over from the beginning and move in a new direction. The Steampunk movement, for example, takes a concept from centuries ago (a reliance on steam power) and then is based in the idea of what would happen if we still used steam power today. The clothing, activities, and games by those in them Steampunk movement have both a retro and futuristic feel. The next creative contribution category is perhaps the most radical. It is *reinitiation*, in which the creator takes the field to a completely new starting point and progresses from there. Many modern and contemporary artists fall into this category. Examples include the impressionist painters, who were the first step in the modern art movement; Auguste Rodin, known as the father of modern sculpture; and Wassily Kandinsky, arguably the first painter to create a nonobjective abstract work of art. Finally there is *integration,* in which the creative contributor takes two seemingly

unrelated fields and merges them into a new product. A good example of this concept is smartphones, which merge the benefits of phones and personal computers.

Although Big-C creativity is not the kind of creativity that educators will focus on cultivating in the classroom, teachers can still include inspirational examples and biographies of Big-C creators in their lessons to help students develop their understanding of the full trajectory of creativity (e.g., Gardner, 1993) and what it takes to make a breakthrough contribution in the various fields of study (e.g., mathematics, science, literature, and so on). In the following, we do not explicitly highlight developmental research on creativity but focus on levels of creativity that can be fostered within a classroom context. Readers interested in the development of creativity can consult Feldman (1999), Feist (2006), Feist and Barron (2003), and Russ and Fiorelli (2010).

6.3.2 Little-c Approaches

In addition to Big-C creativity, most people recognize that there are many everyday forms of creativity. This could be a creative lesson plan developed by a teacher or students' unique way of solving a math problem. Creativity researchers have recognized that such creative activities – conducted everyday by laypersons or individuals who would not necessarily be considered experts or luminaries – should still be considered creative (e.g., Richards, 2007; Richards, Kinney, Benet, & Merzel, 1988). This level of creativity is referred to as "little-c" creativity.

Little-c creativity is an attainable goal for a classroom context. Not only can teachers strive to make their own little-c creative contributions in how they design and teach lessons, but they can also help students develop their ability to make little-c contributions as part of their learning of academic topics. A science teacher, for instance, can help encourage students to develop original and task-appropriate projects of inquiry. A language arts teacher can support students in developing their own unique and meaningful stories, poems, and plays. A social studies teacher can help students develop a novel and meaningful oral history project of their community. A math teacher can encourage students to come up with new and different ways of solving a math problem. The key to little-c creativity is the combination of originality and appropriateness in the context of a particular assignment or learning task. This conception fits with the standard definition of creativity, which emphasizes the combination of novelty and usefulness as defined in a particular sociocultural context (Plucker, Beghetto, & Dow, 2004).

This definition highlights the important role that the sociocultural context (e.g., sixth grade social studies class vs. an exhibit) plays in determining what will be considered novel and useful. If a group of students present their oral history project to their classmates, they would not expect their peers and teachers to begin the critique by comparing the project to a display at the Natural History Museum. Indeed, to reach that level of accomplishment usually takes 10 or more years of developing one's expertise (Hayes, 1989). The peers and teachers would instead evaluate the

presentation by determining whether there was something new and original about it and to make sure that the presentation followed the project guidelines outlined by the teachers.

In this way, little-c creativity is a learning objective that can be easily incorporated in the evaluative criteria of most any academic project or assignment (Beghetto & Kaufman, 2009). A key question for educators, however, is how to develop students' creative potential into little-c creative contributions (see Beghetto et al., 2015, for detailed, specific examples of lessons and best practices). Indeed, students' initial ideas and insights about a project likely will not be judged to be novel or useful. As such, standard (little-c) definitions of creativity are not applicable (because the standard definition relies on external judgments of novelty and usefulness).

Importantly, however, students' early insights can still represent ideas that are personally new and meaningful to *them*. Even though their ideas likely will not represent anything new to the field, the very process of learning a field presents many opportunities for students to experience creative insights and interpretation. In other words, a student's first attempt at writing a short story might not be considered creative by other people who might read the story – it might be derivative of a recent movie or television show or it might be filled with spelling and grammatical mistakes that make it difficult to understand. If the story was being evaluated for creativity by an outside judge, it might be given a very low score. But the experience of writing the story may be a creative experience for the student. Perhaps the student loves writing the story and experiences very high levels of intrinsic motivation (which is strongly associated with creativity, e.g., Amabile, 1996; Hennessey, 2010). Or the student might have new thoughts and ideas that lead them to seek out new experiences or take risks.

6.3.3 Beyond Big and Little

The distinction between Big-C and little-c creativity has been useful for helping to clarify different levels of creative magnitude. However, these two categories do not sufficiently capture the various levels of creative magnitude, which has tremendous implications for all aspects of the creative process. The level of magnitude, for example, strongly impacts basic assessment, as we discuss later in the chapter. Further, the ways to improve creativity are also dependent on the level. For example, the way that a mentor may offer feedback or instruction may vary based upon the level of the four Cs.

How might we think about the creative interpretations made by students as they learn something new? What if these insights and interpretations are new and meaningful only to the student? Should those insights and interpretations still be considered creative? If all we have is Big-C and little-c creativity, then such interpretations would not be considered creative because they do not fit either level of creativity. It seems problematic, however, to simply dismiss such experiences as uncreative.

Indeed, subjective forms of creativity (Stein, 1953) are often given short shrift in Big-C and little-c conceptions of creativity. Runco (2004) argues that this is part of a larger trend in creativity research in which scholars value the creative product so much that they end up overlooking creative potential.

Consider another example. How about people who are extremely creative and accomplished, but not at the Big-C level – should they be grouped with little-c creators? If so, then little-c becomes too broad a category. Big-C and little-c conceptions of creativity simply are not nuanced enough to capture the complex nature of creativity. This gap was a driving force behind the Four-C Model of Creativity (Kaufman & Beghetto, 2009, 2013a, 2013b; Beghetto & Kaufman, 2007), which proposed two additional categories: "Pro-c" and "mini-c." Pro-c creativity is the level below Big-C creativity, but above little-c creativity. For example, if we consider the field of architecture, a Big-C architect might be Frank Lloyd Wright. A Pro-c architect might be someone like Eugene Pandala, an award-winning creator who may very well be considered Big-C in the future.

The lowest level of the four-C hierarchy is mini-c, which is characterized as subjective experiences of creativity that occur naturally during the learning process. Mini-c consists of new insights and meaningful self-discovery, which may not be deemed novel and useful by an external judge. Continuing the example, a beginning architecture student would be at the mini-c level, where an advanced student or beginning professional who has demonstrated creativity that is recognizable to others could be considered little-c.

One way of viewing the hierarchy of the Four-C Model is as a series of developmental stages. With appropriate feedback, mini-c can develop into little-c. With enough practice, little-c can become Pro-c. And given enough cultural impact and time to provide context, Pro-c may eventually be viewed as Big-C.

Given that mini-c is most applicable to classroom settings, we elaborate on the concept in the section that follows.

6.3.4 Mini-c Creativity in the Classroom

Mini-c creativity offers teachers an important way of thinking about student creativity. If teachers only think about larger-C conceptions of creativity or creativity as representing unconstrained originality, then it can be difficult for them to understand how creativity could play a role in their classroom. Indeed, overly narrow conceptions of creativity can serve as the basis for negative stereotypes and perceptions about creativity (Plucker et al., 2004).

Creativity researchers working across cultures have documented various negative beliefs and perceptions about creative students. For instance, educators sometimes associate creativity with nonconformity, impulsivity, and disruptive behavior. Consequently, creative students may be valued less than bright students (e.g., Baudson & Preckel, 2013; Dawson, 1997; Scott, 1999). This is not to say that all teachers devalue creativity. Researchers have also documented examples of teachers

who feel more favorably about creative students (e.g., Runco, Johnson, & Baer, 1993). Importantly, however, teachers often are not clear on what creativity means (Aljughaiman & Mowrer-Reynolds, 2005) and their conceptions can be at odds with how researchers have defined creativity.

Westby and Dawson (1995), for example, found that teachers who reported liking creative students tended to define creativity with adjectives such as "well behaved" or "conforming." These same teachers were then given descriptors that were more typically used to describe creative people, such as "impulsive" or "tries to do what others call impossible" – and their perceptions were then reported to be more negative. In another study, Diakidoy and Phtiaka (2002) discovered that teachers associated creativity primarily with the arts and did not associate "knowledge" as a meaningful component of creativity (see also Seo, Lee, & Kim, 2005). Further, de Souza Fleith (2000) found that although teachers articulated how their attitudes might impact student creativity, they did not consider concepts such as self-evaluation, rewards, or intrinsic motivation as being related to creativity.

Negative or misinformed perceptions of creativity can transcend cultures. For example, Tan (2003) reported that student teachers in Singapore favored students who had pleasant dispositions (e.g., kind, friendly) over students who were more creative and risk taking. Chan and Chan (1999) found that Chinese teachers associated socially undesirable traits with student creativity, arguing that in Chinese cultures, nonconforming or expressive behavior can be interpreted as arrogant or rebellious. Similar findings have been found in Turkish teachers (Güncer & Oral, 1993). Hence, regardless of cultural context, negative stereotypes that are not rooted in truth can diminish creativity in the classroom.

Creativity scholars have long recognized that creativity plays a key role in learning. Guilford (1950), for instance, asserted "a creative act is an instance of learning" and that any "comprehensive learning theory" must take creativity into account (p. 446). Jean Piaget, the eminent learning theorist, also recognized this connection. His "constructivist" view of learning represents a deep, conceptual recognition of the role that creativity plays in process of learning. In fact, he titled his book that outlines his views on learning and written for general public, *To understand is to invent* (Piaget, 1973). If teachers broaden their conceptions to include mini-c creativity, then they will be in a better position to understand how creativity can play a meaningful role in their classroom. Rather than creativity being viewed as something outside of the scope of classroom learning, mini-c stresses that creative insights and interpretations are present in students' everyday learning of core curricular topics.

Consider two students who are asked to solve a math problem. One student simply memorized and reproduced the procedure demonstrated by the teacher to solve this type of problem. The student can produce correct answers, but doesn't really understand why. Another student had a new and personally meaningful insight about how to solve such problems based on her prior experiences running a lemonade stand. She too can produce correct answers. Her approach is a bit unconventional, but she has a clear understanding of why her method works. Although both

students receive the same grade on their assignment, student two likely has a deeper understanding of this type of problem.

Of course, not all learning experiences are as clear-cut as this example. Still, if students have an opportunity to make new and personally meaningful (mini-c) insights when learning, then they are more likely to develop a deeper understanding. By "deeper understanding," we simply mean that it is more personally meaningful. When learning is personally meaningful, students will likely experience intrinsic motivation (in other words, they will engage in the material because of their own inherent interest as opposed to an extrinsic reason such as getting a better grade). Intrinsic motivation has been linked to both, increased academic performance (Deci & Ryan, 2000, Ryan & Deci, 2000) and creativity (Amabile, 1996; Hennessey, 2010). In this way, mini-c learning experiences can serve as a motivational driver for the further development of their academic and creative capacity. Mini-c creativity thereby reframes creativity in a more positive and feasible light for teachers. When teachers recognize this, they will be in a better position to help students' develop their mini-c insights into little-c contributions (Beghetto & Kaufman, 2010).

6.4 Creativity and Achievement

Is creativity related to student achievement? There are at least two ways to approach this question. One way is examine whether there is a relationship between student creativity and student achievement. The other way is to explore whether there is a connection between creative teaching and student achievement. Each will be discussed in turn.

6.4.1 Relation of Creativity and Achievement

Researchers have long been curious about the link between creativity and achievement. Much of the empirical work to date has demonstrated a relationship, but the specific nature and magnitude of this link has been inconsistent and equivocal. Some of the earliest research that examined this link demonstrated relatively strong connections (e.g., Cline, Richards, & Needham, 1963; Getzels & Jackson, 1962), whereas other researchers (e.g., Cicirelli, 1965; Silvia, 2008) have found weaker associations. In more recent years, researchers have found that creativity serves as a significant predictor of academic achievement, but often not as strong a predictor as other traits such as reasoning ability (Freund & Holling, 2008) or cognitive style (Niaz, Saud de Nunez, & Ruiz de Pineda, 2000). Others have found no relationship (or even a negative connection) between creativity and achievement (Toth & Baker, 1990).

Researchers have also examined how creativity and related factors (e.g., divergent thinking) are connected with grade-point averages (Freund & Holling, 2008; Vock, Preckel, & Holling, 2011). Grigorenko et al. (2009), for instance, found a link between creativity tasks and GPA in students who were enrolled in a highly competitive private school (Choate Rosemary Hall). Choate students who performed well on creative writing and creative science tasks tended to have higher GPAs. Moreover, the creative science task significantly predicted Choate students' first year GPA. Other researchers have found additional variations in the relationship between creativity and GPA. Altman (1999), for example, found bimodal distribution in the correlations between a composite measure of divergent thinking and GPA; creativity was most strongly associated with early course grades and advanced course grades. Another study of creativity and achievement, which focused on lower economic status of Spanish secondary students (Ai, 1999), demonstrated that the link between divergent thinking and achievement varied by gender, subject areas, and the type of creativity measure used.

Finally, a recent study by Gralewksi and Karwowski (2012) provides an even more complicated picture of the relationship between divergent thinking and drawing production and school grades (GPA) of high school students in Poland. They found that in some schools creative abilities were positively correlated with school grades, whereas in many other schools, there was no and even a negative relation between creative ability and school grades (see also Freund & Holling, 2008, for a similar finding). Interestingly, in the schools where creative ability played an important role in GPA, the influence of student intelligence was relatively small. Creativity also tended to play a more important role in larger schools, which were located in larger cities. In summarizing these findings, the authors noted that creativity tended not to be reflected in the GPA in the schools they studied. A positive relationship between creativity and school grades was "more of the exception than the rule" (p. 206).

Other studies have examined the relationship between creativity and different facets of achievement. Some researchers, for example, have examined the relationship between students' creative self-beliefs and teachers' ratings of math and science understanding. Beghetto and Baxter (2012), for example, found that whereas students' creative self-beliefs were indirectly related to science understanding, creative math self-beliefs were directly related to math understanding.

A related area of study has involved researchers examining what cognitive factors might predict creative achievement. Kim (2008), for instance, conducted a meta-analysis examining whether IQ or divergent thinking was a stronger predictor of creative achievement. Aside from musical achievement, divergent thinking tests served as the strongest predicator of various types of creative achievement.

What might be said about these various results? One way to interpret these findings is to acknowledge that although there is evidence of a link between creativity and achievement, there seem to be a variety of factors that influence this relationship. One of the most obvious is the use of different measures and assessment methods (see next chapter). Measures of creativity that rely more on divergent thought

have yielded modest and sometimes mixed results. Similarly, the use of domain general measures of creativity may fail to be related to more domain-specific measures of achievement (Baer, 2012). The learning environment (Amabile & Pilemer, 2012; Beghetto & Kaufman, 2014), cultural milieu (Niu & Kaufman, 2013), and instructional style (Schacter et al. 2006) may also influence the relationship between creativity and academic achievement.

6.4.2 The Role of Teachers

Teachers play a key role in supporting student achievement (Schacter et al., 2006). A key question, however, is whether teaching for creativity is related to students' academic achievement. The empirical work that has directly addressed this question is limited. One particularly insightful example is a study conducted by Schacter et al. (2006). Schacter and his team developed a creative teaching framework and used it to score the instructional practices of 48 upper-elementary teachers.

For each teacher, the researchers observed their teaching eight times across the school year. The researchers scored the frequency and quality of teachers' creative instructional behaviors. Creative teaching behaviors included a variety of factors, such as explicitly teaching creative thinking strategies, providing opportunities for choice and discovery, encouraging intrinsic motivation, establishing a learning environment conductive to creativity, and providing opportunities for imaginative thinking.

Schacter and his team found that teachers who used creativity-supportive strategies produced substantial achievement gains in their students (as measured by pre-/post-standardized tests of math, reading, and language). Unfortunately, the majority of teachers they observed did not implement any teaching strategies that fostered student creativity. Moreover, classrooms with high proportions of low performing and minority students were significantly less likely to have teachers who taught for creativity. This study indicates that although teaching for creativity was associated with student achievement, teachers need the freedom, training, and support necessary to infuse these approaches into their everyday teaching.

6.5 Assessment *of* and *for* Creativity

Creativity assessment in schools is typically done for research or for making placement decisions in gifted and talented programs. It is worth noting that most approaches to giftedness include creativity as a central part of the definition. The federal government proposed a multifaceted definition in the early 1970s that highlighted six areas that can highlight giftedness: general intellectual ability, specific academic aptitude, creative or productive thinking, leadership ability, visual and

performing arts, and psychomotor ability (Marland, 1972). This conception has continued to be very influential and is still used by many school districts (e.g., Callahan, Hunsaker, Adams, Moore, & Bland, 1995). Hence, gifted admissions represent one of the primary applications for creativity assessment (Kaufman, Plucker, & Russell, 2012).

Researchers have developed and used a wide variety of assessment methods and techniques. By far, the most frequently used assessment in school settings is the Torrance Tests of Creative Thinking (TTCT). Indeed, the TTCT is the most popular test of divergent thinking used in creativity studies (Plucker & Makel, 2010). The TTCT is an assessment of divergent thinking, which is an aspect of creativity – but not its sole component. Other measures may ask the students about their own creativity to get a self-assessment (e.g., Beghetto, 2006; Beghetto, Kaufman, & Baxter, 2011; Kaufman, 2012) or may ask teachers for their assessment of student creativity (Renzulli, Smith, White, Callahan, & Hartman, 1977) or utilize subject matter experts (Amabile, 1996; Kaufman & Baer, 2012).

The TTCT and these other measures used by researchers are, quite frankly, little use to teachers. They represent assessments *of* creativity or components of creativity. Much more useful for teachers are assessments *for* creativity. Specifically, assessments that help teachers support students' mini-c creative potential transform into little-c contributions. In the following section, we describe task-specific feedback as one way of assessment *for* creativity and fostering of students' creative metacognitions as an intervention to enhance their creativity.

Simple things can make a big difference. In math, for example, teachers can require students to not only find the correct response but also come up with as many ways to solve the problem as they can or even come up with their own problems. Some teachers already do this in their classroom. The key is acknowledging that such practices support creativity and then incorporating these practices more systematically into one's everyday teaching and assessments (Beghetto, 2013).

How students interpret their successes and failures plays a key role in determining whether they will persist in the face of challenges and, in turn, develop their academic and creative competence (Beghetto, 2013; Dweck, 2000). If students feel shamed by their failures and that there is no hope of improvement, then they are less likely to continue putting forth effort in developing their competence (Beghetto, 2014). Teachers can help students interpret the everyday challenges, setbacks, and failures in a more positive light. Indeed, one of the best ways that teachers can support students' academic and creative development is through task-specific feedback (Beghetto, 2007, 2013; Beghetto & Kaufman, 2007). How might this feedback look in the classroom?

Beghetto (2007) discussed how teachers can use supportive feedback to encourage movement from mini-c interpretations to little-c expressions by (a) taking the time to hear and attempt to understand how students are interpreting what they are learning; (b) helping students recognize when their contributions are not making sense given the domain constraints, conventions, and standards of a particular activity or task; and (c) providing multiple opportunities for students to practice developing

the skills of a particular domain or task. Put simply, supportive feedback emphasizes the creative strengths of students (what they can already do well) and also notes creative limitations (where they need to keep working). Teachers may find our *Goldilocks Principle* of feedback (Beghetto & Kaufman, 2007) useful in serving as a reminder for how to provide balanced, supportive feedback. Specifically, this principle argues for the importance of giving feedback that is not too harsh (stifling student motivation) yet not too gentle (with little attention to real-world standards).

An example of balanced feedback may help clarify. Imagine a student, Sophia, who submits a short story to the high school literary magazine. The content of the story is promising, but needs work. If she is judged too harshly, then she may come to believe that she is not a creative writer and stop writing even for her own enjoyment (see Beghetto, 2014). If she is over praised, then she will be ill equipped when she receives honest, real-world feedback; consider the candidates on *American Idol* or other shows that are wildly overmatched and humiliated when their poor singing gets a harsh and public critique. Many of these singers auditioned for the show when family and friends, meaning well, encouraged and praised poor performance. Although well-intentioned, this practice can lead to people having poor self-insight into their creative abilities (Kaufman, Evans, & Baer, 2010).

Sophia may thus experience a public embarrassment when the story is published (or rejected) and not understand why. Empty praise cheats students from receiving the kind of demanding feedback necessary for creative growth. If, however, she receives more detailed and demanding feedback, which highlights how specifically she can improve, then she will be in a better position to develop her creative story-writing skills. She will also have a more solid self-perception of her creative abilities and better able to develop her mini-c efforts into little-c (and perhaps someday Pro-c) contributions.

The task-specific feedback and fostering practice described in this section often represent moment-to-moment instructional decisions. These decision points are dynamic and sometimes difficult to predict. This is one reason why teaching for creativity is sometimes described as an improvisational performance (Beghetto & Kaufman, 2011) rather than a highly predictable sequence of steps. Consequently, this presents challenges both to teachers who want to learn how to teach in this way and to researchers who may be interested in understanding the factors underlying these practices.

For teachers, habits of practice do not develop overnight. Fortunately, however, they can be developed with minor, everyday adjustments made to existing practices. For researchers, the kinds of studies and interventions that seem to be of greatest use would be those that would be the most costly and intensive (e.g., observation and intervention studies conducted over time). This is because more intensive studies may help preserve the complexity and dynamic context of actual classroom instruction. That said, research using smaller numbers of observation or simulations may also yield helpful insights and is worth exploring.

6.6 Creative Metacognition

We feel that one of the best ways to support the development of students' creative competence is to help them take charge of their own creativity. Knowing when and when not to be creative is a skill that more accomplished creators have mastered. Teachers play a key role in helping students develop this skill. Specifically, teachers need to teach students to know when (and when not) to be creative. This knowledge, called creative metacognition (CMC) (Kaufman & Beghetto, 2013b), refers to "a combination of creative self-knowledge (knowing one's own creative strengths and limitations, both within a domain and as a general trait) and contextual knowledge (knowing when, where, how, and why to be creative)" (p. 160). Like other forms of metacognitive knowledge (Flavell, 1979; Pintrich, Wolters, & Baxter, 2000), CMC is a particular form of cognition that helps people monitor and develop their creative competence. There may be times, for example, when there are risks involved in being creative – from the student's perspective or the teacher's perspective. A student being creative "at the wrong time" can waste time, disturb or pester other classmates, or potentially be embarrassed or ignored. We are not even referring to the extreme levels of potential "malevolent creativity" (Cropley, Cropley, Kaufman, & Runco, 2010; Cropley, Kaufman, & Cropley, 2008). It is important to be able to harness the potential power of creativity so that it can have the best possible impact. Teachers can help their students develop the knowledge necessary to "read the situation" and determine where the context represents a good time to publicly express creativity.

Within the K-12 classroom, teachers can help students understand the situations that are most likely to be receptive to creative expression. Teachers might support CMC when they give instructions for an assignment and when they give feedback, thereby explaining the specific limits and expectations that might be present for a particular activity. Teachers can help students distinguish not only when original thinking or intellectual risk taking would be particularly helpful on a specific assignment but also help explore the best ways that it could be expressed (e.g., "I want you to first solve this algebraic proof in the way that I have showed you, to demonstrate that you understand each of the properties. Once you have demonstrated that you understand the basics, I am happy for you to experiment.").

6.7 Concluding Summary and Resources

Our aim in this chapter was to help promote an understanding of creativity in schools and classrooms. To this end we reviewed definitions and conceptions of creativity. We focused on definitions and conceptions that we feel are most appropriate for K-12 settings. Specifically, we discuss the Four-C Model of Creativity and implications for this model. We also discussed the relationship between creativity and achievement, creativity assessment, and the importance of cultivating creative metacognition.

In what follows, we briefly summarize a few key insights and provide a list of recent resources for educators interested in cultivating students' creative potential in the classroom.

6.7.1 Summary of Key Insights

- **What is creativity**? Creativity involves a combination of originality and appropriateness. Judgments of creativity also occur in a particular context.
- **How to think about creativity?** Researchers have categorized creativity in many ways. The Four-C Model highlights a developmental progression that includes four levels of creativity: mini-c (subjective), little-c (everyday), Pro-c (professional), and Big-C (legendary) creativity.
- **What is classroom creativity?** In most cases, the levels of creativity of most concern to teachers are mini-c and little-c creativity. Teachers can support the development of students' mini-c insights into little-c contributions by providing supportive feedback.
- **Is there a link between creativity and learning?** Creativity researchers and several prominent learning theorists have long asserted that creativity and learning are linked. Research has also demonstrated a link between creativity and achievement. The findings, however, are somewhat equivocal. Additional research is needed to better understand the nature of this relationship in school and classroom contexts.
- **What can teachers do to support creativity?** There are several things that teachers can do to support creativity. Many of our suggestions described herein simply require increased awareness and making slight adjustments to one's existing instructional and assessment practices. Honest and supportive feedback is one of the best ways to encourage and support student creativity. Once teachers have a working understanding of creativity, they can then actively encourage it from and demonstrate it for their students.

6.7.2 Creativity Resources for Educators

- Baer, J., & Kaufman, J. C. (2013). *Being creative inside and outside the classroom.* Rotterdam: Sense Publishers.
- Beghetto, R. A. (2013). *Killing ideas softly? The promise and perils of creativity in the classroom.* Charlotte, NC: Information Age Publishing.
- Beghetto, R. A., & Kaufman, J. C. (Eds.). (2010). *Nurturing creativity in the classroom.* New York, NY: Cambridge University Press.
- Beghetto, R. A., Kaufman, J. C., & Baer, J. (in press). *Creativity and the common core.* New York: Teachers College Press.

- Craft, A. (2011). *Creativity and education futures: Learning in a digital age.* Stoke-on-Trent, UK: Trentham.
- Daniels, S., & Peters, D. A. (2013). *Raising creative kids*. Scottsdale, AZ: Great Potential Press, Inc.
- Gregerson, M., Kaufman, J. C., & Snyder, H. (Eds.). (2013). *Teaching creatively and teaching creativity.* New York, NY: Springer Science.
- Kaufman, J. C. (2009). *Creativity 101*. New York: Springer.
- Kaufman, S. B. (2013). *Ungifted*. New York: Basic Books.
- Piirto, J. (2011). *Creativity for 21st century skills: How to embed creativity into the curriculum*. Rotterdam: Sense Publishers.
- Piirto, J. (Ed.). (2013). *Organic creativity in the classroom.* Scottsdale, AZ: Great Potential Press.
- Plucker, J. A., & Callahan, C. M. (Eds.). (2014). *Critical issues and practices in gifted education: What the research says* (2nd Ed.). Waco, TX: Prufrock Press.
- Saracho, O. N. (Ed.). (2012). *Contemporary perspectives on research in creativity in early childhood education.* Charlotte, NC: Information Age Publishing.
- Sawyer, R. K. (Ed.). (2011). *Structure and improvisation in creative teaching.* New York, NY: Cambridge University Press.
- Sawyer, R. K. (2012). *Explaining creativity: The science of human innovation.* Oxford, UK: Oxford University Press.
- Tan, A. G. (Ed.). (2007). *Creativity: A handbook for teachers*. Singapore: World Scientific.

References

Ai, X. (1999). Creativity and academic achievement: An investigation of gender differences. *Creativity Research Journal, 12*, 329–337.

Aljughaiman, A., & Mowrer-Reynolds, E. (2005). Teachers' conceptions of creativity and creative students. *Journal of Creative Behaviour, 39*, 17–34. doi:10.1002/j.2162-6057.2005.tb01247.x.

Altman, W. S. (1999). *Creativity and academic success*. Unpublished dissertation, Cornell University, Ithaca, NY.

Amabile, T. M. (1996). *Creativity in context: Update to the social psychology of creativity*. Boulder, CO: Westview.

Amabile, T. M., & Pillemer, J. (2012). Perspectives on the social psychology of creativity. *Journal of Creative Behavior, 46*, 3–15.

Baer, J. (2012). Domain specificity and the limits of creativity theory. *Journal of Creative Behaviour, 46*, 16–29. doi:10.1002/jocb.002.

Baer, J., & Kaufman, J. C. (2006). Creativity in English-speaking countries. In J. C. Kaufman & R. J. Sternberg (Eds.), *International handbook of creativity* (pp. 10–38). New York: Cambridge University Press.

Barron, F. (1955). The disposition toward originality. *Journal of Abnormal and Social Psychology, 51*, 478–485.

Baudson, T. G., & Preckel, F. (2013). Teachers' implicit personality theories about the gifted: An experimental approach. *School Psychology Quarterly, 28*(1), 37–46. doi:10.1037/spq0000011.

Beghetto, R. A. (2006). Creative self-efficacy: Correlates in middle and secondary students. *Creativity Research Journal, 18*, 447–457. doi:10.1207/s15326934crj1804_4.

Beghetto, R. A. (2007). Ideational code-switching: Walking the talk about supporting student creativity in the classroom. *Roeper Review, 29*, 265–270. doi:10.1080/02783190709554421.

Beghetto, R. A. (2013). *Killing ideas softly? The promise and perils of creativity in the classroom*. Charlotte, NC: Information Age Publishing.

Beghetto, R. A. (2014). Creative mortification: An initial exploration. *Psychology of Aesthetics, Creativity, and the Arts, 8*(3), 266–276.

Beghetto, R. A., & Baxter, J. A. (2012). Exploring student beliefs and understanding in elementary science and mathematics. *Journal of Research in Science Teaching, 49*, 942–960. doi:10.1002/tea.21018.

Beghetto, R. A., & Kaufman, J. C. (2007). Toward a broader conception of creativity: A case for "mini-c" creativity. *Psychology of Aesthetics, Creativity, and the Arts, 1*, 13–79. doi:10.1037/1931-3896.1.2.73.

Beghetto, R. A., & Kaufman, J. C. (2009). Intellectual estuaries: Connecting learning and creativity in programs of advanced academics. *Journal of Advanced Academics, 20*, 296–324. doi:10.1177/1932202X0902000205.

Beghetto, R. A., & Kaufman, J. C. (Eds.). (2010). *Nurturing creativity in the classroom*. New York: Cambridge University Press.

Beghetto, R. A., & Kaufman, J. C. (2011). Teaching for creativity with disciplined improvisation. In R. K. Sawyer (Ed.), *Structure and improvisation in creative teaching*. Cambridge: Cambridge University Press.

Beghetto, R. A., & Kaufman, J. C. (2013). Fundamentals of creativity. *Educational Leadership, 70*, 10–15.

Beghetto, R. A., & Kaufman, J. C. (2014). Classroom contexts for creativity. *High Ability Studies, 25*, 53–69. doi:10.1080/13598139.2014.905247.

Beghetto, R. A., Kaufman, J. C., & Baer, J. (2015). *Creativity and the common core*. New York: Teachers College Press.

Beghetto, R. A., Kaufman, J. C., & Baxter, J. (2011). Answering the unexpected questions: Exploring the relationship between students' creative self-efficacy and teacher ratings of creativity. *Psychology of Aesthetics, Creativity, and the Arts, 5*, 342–349. doi:10.1037/a0022834.

Beghetto, R. A., & Plucker, J. A. (2006). The relationship among schooling, learning, and creativity: "All roads lead to creativity" or "You can't get there from here?". In J. C. Kaufman & J. Baer (Eds.), *Creativity and reason in cognitive development* (pp. 316–332). Cambridge, UK: Cambridge University Press. doi:10.1017/CB09780511606915.019.

Bonett, D. G. (2012). Replication-extension studies. *Current Directions in Psychological Science, 21*, 409–412. doi:10.1177/0963721412459512.

Callahan, C. M., Hunsaker, S. L., Adams, C. M., Moore, S. D., & Bland, L. C. (1995). *Instruments used in the identification of gifted and talented students (Report No. RM-95130)*. Charlottesville, VA: National Research Center on the Gifted and Talented.

Chan, D. W., & Chan, L. K. (1999). Implicit theories of creativity: Teachers' perception of student characteristics in Hong Kong. *Creativity Research Journal, 12*, 185–195. doi:10.1207/s15326934crj1203_3.

Cicirelli, V. G. (1965). Form of relationship between creativity, IQ and academic achievement. *Journal of Educational Psychology, 56*, 303–308. doi:10.1037/h0022792.

Cline, V. B., Richards, J. M., Jr., & Needham, W. E. (1963). Creativity tests and achievement in high school science. *Journal of Applied Psychology, 47*, 184–189. doi:10.1037/h0041145.

Cropley, D. H., Cropley, A. J., Kaufman, J. C., & Runco, M. A. (Eds.). (2010). *The dark side of creativity*. New York: Cambridge University Press.

Cropley, D. H., Kaufman, J. C., & Cropley, A. J. (2008). Malevolent creativity: A functional model of creativity in terrorism and crime. *Creativity Research Journal, 20*, 105–115. doi:10.1080/10400410802059424.

Csikszentmihalyi, M. (1999). Implications of a systems perspective for the study of creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 313–335). New York: Cambridge University Press.

Dawson, V. L. (1997). In search of the wild bohemian: Challenges in the identification of the creatively gifted. *Roeper Review, 19*, 148–152. doi:10.1080/02783199709553811.
de Souza Fleith, D. (2000). Teacher and student perceptions of creativity in the classroom environment. *Roeper Review, 22*, 148–153. doi:10.1080/02783190009554022.
Deci, E. L., & Ryan, R. M. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*, 68–78. doi:10.1037/0003-066X.55.1.68.
Diakidoy, I. N., & Phtiaka, H. (2002). Teachers' beliefs about creativity. In S. P. Shohov (Ed.), *Advances in psychology research* (Vol. 15, pp. 173–188). Hauppauge, NY: Nova Science.
Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality and development*. Philadelphia: Taylor & Francis.
Feist, G. J. (2006). How development and personality influence scientific thought, interest, and achievement. *Review of General Psychology, 10*, 163–182. doi:10.1037/1089-2680.10.2.163.
Feist, G. J., & Barron, F. X. (2003). Predicting creativity from early to late adulthood: Intellect, potential and personality. *Journal of Research in Personality, 37*, 62–88. doi:10.1016/S0092-6566(02)00536-6.
Feldman, D. H. (1999). The development of creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 169–186). Cambridge: Cambridge University Press.
Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906–911. doi:10.1037/0003-066X.34.10.906.
Freund, P. A., & Holling, H. (2008). Creativity in the classroom: A multilevel analysis investigating the impact of creativity and reasoning ability on scholastic achievement. *Creativity Research Journal, 20*, 309–318. doi:10.1080/10400410802278776.
Gardner, H. (1993). *Creating minds*. New York: Basic Books.
Getzels, J. W., & Jackson, P. W. (1962). *Creativity and intelligence*. New York: Wiley.
Glaveanu, V. P. (2013). Rewriting the language of creativity: The five A's framework. *Review of General Psychology, 17*, 69–81. doi:10.1037/a0029528.
Gralewksi, J., & Karwowski, M. (2012). Creativity and school grades: A case from Poland. *Thinking Skills and Creativity, 7*, 198–208. doi:10.1016/j.tsc.2012.03.002.
Grigorenko, E. L., Jarvin, L., Diffley, R., Goodyear, J., Shanahan, E., & Sternberg, R. J. (2009). Are SATS and GPA enough? A theory-based approach to predicting academic success in secondary school. *Journal of Educational Psychology, 101*, 964–981.
Guilford, J. P. (1950). Creativity. *American Psychologist, 5*, 444–454. doi:10.1037/h0063487.
Güncer, B., & Oral, G. (1993). Relationship between creativity and nonconformity to school discipline as perceived by teachers of Turkish elementary school children, by controlling for their grade and sex. *Journal of Instructional Psychology, 20*, 208–214.
Hayes, J. R. (1989). *The complete problem solver* (2nd ed.). Mahwah, NJ: Erlbaum.
Hennessey, B. A. (2010). Intrinsic motivation and creativity in the classroom: Have we come full circle? In R. A. Beghetto & J. C. Kaufman (Eds.), *Nurturing creativity in the classroom*. New York, NY: Cambridge University Press.
Kaufman, J. C. (2009). *Creativity 101*. New York, NY: Springer.
Kaufman, J. C. (2012). Counting the muses: Development of the Kaufman-Domains of creativity scale (K-DOCS). *Psychology of Aesthetics, Creativity, and the Arts, 6*, 298–308. doi:10.1037/a0029751.
Kaufman, J. C. (Ed.). (2014). *Creativity and mental illness*. New York: Cambridge University Press.
Kaufman, J. C., & Baer, J. (2004). Hawking's haiku, Madonna's math. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 3–20). Washington, DC: American Psychological Association.
Kaufman, J. C., & Baer, J. (2012). Beyond new and appropriate: Who decides what is creative? *Creativity Research Journal, 24*, 83–91. doi:10.1080/10400419.2012.649237.
Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The Four C model of creativity. *Review of General Psychology, 13*, 1–12. doi:10.1037/a0013688.
Kaufman, J. C., & Beghetto, R. A. (2013a). Do people recognize the Four Cs? Examining layperson conceptions of creativity. *Psychology of Aesthetics, Creativity, and the Arts, 7*, 229–236. doi:10.1037/a0033295.

Kaufman, J. C., & Beghetto, R. A. (2013b). In praise of Clark Kent: Creative metacognition and the importance of teaching kids when (not) to be creative. *Roeper Review, 35*, 155–165. doi:10.1080/02783193.2013.799413.

Kaufman, J. C., & Beghetto, R. A. (2014). Creativity in the schools: Renewed interest and promising new directions. In M. J. Furlong, R. Gilman, & E. S. Huebner (Eds.), *Handbook of positive psychology in the schools* (pp. 165–175). New York: Routledge.

Kaufman, J. C., Evans, M. L., & Baer, J. (2010). The American Idol Effect: Are students good judges of their creativity across domains? *Empirical Studies of the Arts, 28*, 3–17. doi:10.2190/EM.28.1.b.

Kaufman, J. C., Plucker, J. A., & Russell, C. M. (2012). Identifying and assessing creativity as a component of giftedness. *Journal of Psychoeducational Assessment, 30*, 60–73. doi:10.1177/0734282911428196.

Kim, K. H. (2005). Can only intelligent people be creative? *Journal of Secondary Gifted Education, 16*, 57–66. doi:10.4219/jsge-2005-473.

Kim, K. H. (2008). Meta-analyses of the relationship of creative achievement to both IQ and divergent thinking test scores. *Journal of Creative Behavior, 42*, 106–130.

Marland, S. P. (1972). *Education of the gifted and talented: Report to the Congress of the United States by the U.S. Commissioner of Education*. Washington, DC: Department of Health, Education and Welfare.

Niaz, M., Nunez, G. S., & Pineda, I. R. (2000). Academic performance of high school students as a function of mental capacity, cognitive style, mobility-fixity dimension, and creativity. *Journal of Creative Behavior, 34*, 18–29.

Niu, W., & Kaufman, J. C. (2013). Creativity of Chinese and American people: A synthetic analysis. *Journal of Creative Behavior, 47*, 77–87.

Piaget, J. (1973). *To understand is to invent*. New York: Grossman.

Pintrich, P., Wolters, C., & Baxter, G. (2000). Assessing metacognition and self–regulated learning. In G. Schraw (Ed.), *Metacognitive assessment*. Lincoln, NE: University of Nebraska Press.

Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? *Educational Psychologist, 39*, 83–96. doi:10.1207/s15326985ep3902_1.

Plucker, J. A., & Makel, M. C. (2010). Assessment of creativity. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity* (pp. 48–73). New York, NY: Cambridge University Press.

Renzulli, J. S., Smith, L., White, A., Callahan, C., & Hartman, R. (1977). *Scales for rating the behavioral characteristics of superior students. [manual and ten rating scales]*. Mansfield Center, CT: Creative Learning Press.

Rhodes, M. (1962). An analysis of creativity. *Phi Delta Kappan, 42*, 305–310.

Richards, R. (Ed.). (2007). *Everyday creativity and new views of human nature: Psychological, social, and spiritual perspectives*. Washington, DC: American Psychological Association. doi:10.1037/11595-000.

Richards, R., Kinney, D. K., Benet, M., & Merzel, A. P. C. (1988). Assessing everyday creativity: Characteristics of the lifetime creativity scales and validation with three large samples. *Journal of Personality and Social Psychology, 54*, 476–485. doi:10.1037/0022-3514.54.3.476.

Runco, M. A. (2003). Personal creativity. *Scandinavian Journal of Psychology, 47*, 317–324.

Runco, M. A. (2004). Everyone has creative potential. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 21–30). Washington, DC: American Psychological Association.

Runco, M. A., Johnson, D. J., & Baer, P. K. (1993). Parents' and teachers' implicit theories of children's creativity. *Child Study Journal, 23*, 91–113.

Russ, S., & Fiorelli, J. A. (2010). Developmental approaches to creativity. In J. C. Kaufman & R. J. Sternberg (Eds.), *Cambridge handbook of creativity* (pp. 233–249). New York: Cambridge University Press.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*, 54–67. doi:10.1006/ceps.1999.1020.

Schacter, J., Thum, Y. M., & Zifkin, D. (2006). How much does creative teaching enhance elementary school students' achievement? *Journal of Creative Behavior, 40*, 47–72.

Scott, C. L. (1999). Teachers' biases toward creative children. *Creativity Research Journal, 12*, 321–337. doi:10.1207/s15326934crj1204_10.

Seo, H., Lee, E., & Kim, K. (2005). Korean science teachers' understanding of creativity in gifted education. *Journal of Advanced Academics, 16*, 98–105. doi:10.4219/jsge-2005-476.

Silvia, P. J. (2008). Creativity and intelligence revisited: A latent variable analysis of Wallach and Kogan (1965). *Creativity Research Journal, 20*, 34–39. doi:10.1080/10400410701841807.

Simonton, D. K. (1977). Creative productivity, age, and stress: A biographical time-series analysis of 10 classical composers. *Journal of Personality and Social Psychology, 35*, 791–804. doi:10.1037/0022-3514.35.11.791.

Simonton, D. K. (1990). *Psychology, science, and history: An introduction to historiometry*. New Haven, CT: Yale University Press.

Simonton, D. K. (1997). Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychological Review, 104*, 66–89. doi:10.1037/0033-295X.104.1.66.

Simonton, D. K. (2004). *Creativity in science: Chance, logic, genius, and zeitgeist*. Cambridge, UK: Cambridge University Press.

Simonton, D. K. (2009). Varieties of (scientific) creativity: A hierarchical model of disposition, development, and achievement. *Perspectives on Psychological Science, 4*, 441–452. doi:10.1111/j.1745-6924.2009.01152.x.

Simonton, D. K. (2012). Taking the US Patent Office creativity criteria seriously: A quantitative three-criterion definition and its implications. *Creativity Research Journal, 24*, 97–106. doi:10.1080/10400419.2012.676974.

Stein, M. (1953). Creativity and culture. *Journal of Psychology, 36*, 311–322. doi:10.1080/00223980.1953.9712897.

Sternberg, R. J., & Kaufman, J. C. (2012). When your race is almost run, but you feel you're not yet done: Application of the Propulsion Theory of Creative Contributions to late-career challenges. *Journal of Creative Behavior, 46*, 66–76. doi:10.1002/jocb.005.

Sternberg, R. J., Kaufman, J. C., & Pretz, J. E. (2001). The propulsion model of creative contributions applied to the arts and letters. *Journal of Creative Behavior, 35*, 75–101. doi:10.1002/j.2162-6057.2001.tb01223.x.

Sternberg, R. J., Kaufman, J. C., & Pretz, J. E. (2002). *The creativity conundrum*. Philadelphia, PA: Psychology Press.

Sternberg, R. J., Kaufman, J. C., & Pretz, J. E. (2003). A propulsion model of creative leadership. *The Leadership Quarterly, 14*, 455–473. doi:10.1016/S1048-9843(03)00047-X.

Sternberg, R. J., & Lubart, T. I. (1995). *Defying the crowd: Cultivating creativity in a culture of conformity*. New York, NY: Free Press.

Sternberg, R. J., Pretz, J. E., & Kaufman, J. C. (2003). Types of innovations. In L. Shavinina (Ed.), *The international handbook of innovation* (pp. 158–169). Mahwah, NJ: Lawrence Erlbaum.

Sternberg, R. J., & Zhang, L. F. (2001). *Perspectives on thinking, learning, and cognitive styles*. Mahwah, NJ: Erlbaum.

Tan, A. G. (2003). Student teachers' perceptions of teacher behaviors for fostering creativity: A perspective on the academically low achievers. *Korean Journal of Thinking and Problem Solving, 13*, 59–71.

Toth, L. S., & Baker, S. R. (1990). The relationship of creativity and instructional style preferences to overachievement and underachievement in a sample of public school children. *Journal of Creative Behavior, 24*, 190–198. doi:10.1002/j.2162-6057.1990.tb00540.x.

Vock, M., Preckel, F., & Holling, H. (2011). Mental abilities and school achievement: A test of a mediation hypothesis. *Intelligence, 39*, 357–369.

Westby, E. L., & Dawson, V. L. (1995). Creativity: Asset or burden in the classroom? *Creativity Research Journal, 8*, 1–10. doi:10.1207/s15326934crj0801_1.

Zhang, L. F., Sternberg, R. J., & Rayner, S. (2012). *Handbook of intellectual styles: Preferences in cognition, learning, and thinking*. New York: Springer.

Chapter 7
Conscientiousness in Education: Its Conceptualization, Assessment, and Utility

Lisa E. Kim, Arthur E. Poropat, and Carolyn MacCann

7.1 Introduction

Personality traits are relatively stable patterns of thoughts, feelings, and behaviors across time and situations (Roberts & Jackson, 2008). As outlined in earlier sections of the book, personality is a well-recognized noncognitive construct, which consistently predicts academic performance above and beyond the effects of intelligence (e.g., Noftle & Robins, 2007; Wagerman & Funder, 2007). The five-factor model of personality is the most widely used personality model. This model describes personality in terms of five domains: openness, conscientiousness, extraversion, agreeableness, and neuroticism (Chap. 1).

In education, there has long been a great interest in *whether* and *how* personality is related to academic achievement. Empirical evidence demonstrates that conscientiousness has the strongest association with academic performance of all the Big Five personality traits. A series of meta-analyses reported similar correlations of conscientiousness with academic performance in tertiary education (0.24, 0.23, and 0.23; O'Connor & Paunonen, 2007; Poropat, 2009; Richardson, Abraham, & Bond, 2012). Poropat (2009) also reported conscientiousness/achievement relationship at primary school level (0.28) and secondary school levels of education (0.28 and 0.21,

The views expressed here are those of the authors and do not reflect that of the University of Sydney or Griffith University.

L.E. Kim (✉) • C. MacCann
School of Psychology, The University of Sydney,
Griffith Taylor Building, Sydney, NSW, Australia
e-mail: Lisa.kim@sydney.edu.au; Carolyn.maccann@sydney.edu.au

A.E. Poropat
School of Applied Psychology, Griffith University,
176 Messines Ridge Road, Mount Gravatt, Nathan, QLD 4152, Australia
e-mail: arthur.poropat@griffith.edu.au

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_7

respectively). While it is known the intelligence/achievement relationship declines substantially from primary and tertiary education (possibly due to increasing range restriction; Jensen, 1980), this decline in prediction does *not* occur for conscientiousness. Poropat (2009) found no similar significant decline in correlations with conscientiousness, suggesting that the true relationship of conscientiousness with academic performance may actually *increase* at higher levels of education (if we assume analogous levels of range restriction occur for conscientiousness). As such, differences in conscientiousness between students may constitute a key explanation for differences in educational success or failure.

Chapter 1 described the lexical hypothesis—that trait descriptors which have value to society emerge as single word terms within a natural language. A corollary of the lexical hypothesis is that traits defined by such descriptors relate to socially valued behaviors and outcomes. As academic achievement is a prototypical example of a socially valued outcome, it should therefore be associated with one or more of the traits that emerge from the lexical hypothesis. Empirical evidence supports this view—conscientiousness predicts academic achievement across all stages of education and *all* of the Big Five factors predict academic achievement in the early grades (Poropat, 2009).

There are two mechanisms that account for the link between conscientiousness and academic achievement: (1) internal motivational factors and (2) habitual behaviors that facilitate achievement. Regarding the first mechanism, motivation is a necessary precondition for long-term academic achievement, and the descriptors for conscientiousness encapsulate the will to achieve. Indeed, studies have shown that highly conscientious students are more motivated to strive academically (Chamorro-Premuzic & Furnham, 2005; Digman, 1989; see also Chap. 10). Regarding the second mechanism, certain behaviors are necessary for academic achievement, such as time management, project management, and good study habits. These elements are captured by conscientiousness (MacCann, Fogarty, & Roberts, 2012), which hence contribute to one's level of academic achievement

Given the importance of conscientiousness in the educational domain, this chapter reviews the nature of conscientiousness, its relationship with similar constructs, some issues to be considered when assessing its relationship with academic achievement, traditional and alternative methods to its measurement, and the natural and practiced methods for its enhancement.

7.2 Nature of Conscientiousness

7.2.1 Definition and the Structure of Conscientiousness

The two major theoretical models of personality in current acceptance are arguably the five-factor model (derived from lexical studies, see Chap. 1) and the more recent HEXACO model (Honesty/Humility, Emotionality, eXtraversion, Agreeableness, Conscientiousness, Openness; Lee & Ashton, 2004). The HEXACO is similar to the

Big Five, with the addition of the honesty/humility domain and a slightly differing delineation of the agreeableness/disagreeableness versus emotional stability/neuroticism dimension. Both of these models include conscientiousness as a major domain of personality. As a broad general definition, *conscientiousness describes individual differences in peoples' approach to tasks and task completion*. Highly conscientious people are well organized, self-directed, and goal driven, whereas people low in conscientiousness can be spontaneous, careless, and disorganized. Many personality models also specify an underlying structure of specific facets that aggregate to form an overall level of conscientiousness (e.g., achievement striving, orderliness, self-discipline, and deliberation are all different facets that underlie conscientiousness).

Multiple personality models have included a conscientiousness-like trait since the first such factor was identified by Webb (1915). Although most frequently referred to as conscientiousness, the domain has also been called *prudence* (Hogan & Hogan, 1992), *conformity* or *dependability* (Fiske, 1949; Hogan, 1983), *will to achieve* or *will* (Digman & Takemoto-Chock, 1981; Smith, 1967; Webb, 1915; Wiggins, Blackburn, & Hackman, 1969), and *work* (Peabody & Goldberg, 1989). The underlying facet structure of conscientiousness has also differed from model to model. In the paragraphs below, we outline several models that implicitly or explicitly propose faceted models of conscientiousness. For clarity, Table 7.1 provides a synthesis of how these models delineate conscientiousness into similar facets.

7.2.2 Theoretical Models of Conscientiousness

NEO Personality Inventory—Revised (NEO-PI-R) The NEO-PI-R six-facet model is the most commonly employed model of personality facets (Costa & McCrae, 1992). The six facets of conscientiousness represent (a) competence (self-efficacious), (b) order (personal orderliness), (c) dutifulness (moralistic), (d) achievement striving (will to achieve, striving for excellence), (e) self-discipline (working without hesitation), and (f) deliberation (thinking before acting or speaking; Costa, McCrae, & Dye, 1991). The NEO facet structure was developed in a top-down fashion, using theory rather than empirical evidence. That is, the facets were specified by the investigators in an a priori manner, based on literature reviews (Costa et al., 1991; McCrae & Costa, 1983). Costa et al. (1991) acknowledge that there may be more facets than are included in their model.

Big Five Aspect Scales (BFAS) DeYoung et al. (2007) propose a hierarchical structure of conscientiousness that includes an intermediate level of personality between the specific facets of the NEO-PI-R and the broad domain levels. For conscientiousness, the six NEO-PI-R facets form two aspects of conscientiousness: (1) *industriousness* (composed of achievement striving, competence, and self-discipline) and (2) *orderliness* (composed of deliberation, dutifulness, and order). There is substantial empirical evidence for this distinction. The 6–2–1 model of personality has received support from a major meta-analysis that demonstrated

Table 7.1 Conceptual correspondence between the facets of conscientiousness in nine different models of conscientiousness

BFAS	NEO-PI-R	16PF	HEXACO	AB5C	CPI	HPI	JPI	MPQ
Orderliness	Dutifulness			Dutifulness	Responsibility, socialization	Not autonomous	Responsibility	
	Order		Organization	Orderliness, organization		Organization	Organization	
	Deliberation		Prudence	Cautiousness		Not spontaneous, avoids trouble		Control
Industriousness	Achievement striving		Diligence	Purposefulness, efficiency	Achievement via conformance, well-being			
	Competence			Rationality		Mastery		Harm avoidance
	Self-discipline	Rule consciousness			Self-control	Impulse control, moralistic		
		Perfectionism	Perfectionism	Perfectionism				
					Good impression	Virtuous	Traditional values	Traditionalism

Note: *BFAS* Big Five Aspect Scales (DeYoung, Quilty, & Peterson, 2007), *NEO-PI-R* Neuroticism Extraversion Openness Personality Inventory—Revised (Costa & McCrae, 1992), *16PF* Sixteen Personality Factor Questionnaire (Conn & Rieke, 1994), *HEXACO* Honesty/Humility, Emotionality, eXtraversion, Agreeableness, Conscientiousness Openness inventory (Lee & Ashton, 2004), *AB5C* Abridged Big Five Dimensional Circumplex scale (Hofstee, De Raad, & Goldberg, 1992), *CPI* California Personality Inventory (Gough & Bradley, 1996), *HPI* Hogan Personality Inventory (Hogan & Holland, 2003), *JPI* Jackson Personality Inventory—Revised (Jackson, 1994), *MPQ* Multidimensional Personality Questionnaire (Tellegen, 1990)

differential predictive utility of the aspects for different workplace outcomes (Judge, Rodell, Klinger, Simon, & Crawford, 2013).

Sixteen Personality Factor Questionnaire (16PF) The 16PF measures 16 personality traits (Conn & Rieke, 1994). The 16PF was developed based on three types of data: observational data (L-data), self-report questionnaire (Q-data), and standardized experiment measurements of behaviors (T-data). Although not developed to assess a five-factor structure, second-order factor analysis of the 16 primary factors showed that they fall under five "global" factors closely resembling the Big Five (Chernyshenko, Stark, & Chan, 2001; Conn & Rieke, 1994). Self-control is one of the five global factors, which closely resembles conscientiousness. The primary factors subsumed under self-control differ by researchers. Some researchers include rule consciousness (adherence to rules) and perfectionism (e.g., perfectionistic, self-disciplined, orderly; Roberts, Chernyshenko, Stark, & Goldberg, 2005).

Abridged Big Five Dimensional Circumplex Scale (AB5C) The AB5C is an integration of the Big Five factor structure with the circumplex model to form a five-dimensional circumplex as trait descriptors are argued to be better represented as blends of factors rather than one factor (Hofstee et al., 1992). Therefore, the AB5C represents personality as a blend of varying degrees of strength within pairs of Big Five factors, creating ten two-dimensional circumplexes. For the case of conscientiousness, there are nine facets in this model: conscientiousness (pure conscientiousness), efficiency (high conscientiousness, high extraversion), cautiousness (high conscientiousness, low extraversion), dutifulness (high conscientiousness, high agreeableness), rationality (high conscientiousness, low agreeableness), purposefulness (high conscientiousness, high emotional stability), perfectionism (high conscientiousness, low emotional stability), organization (high conscientiousness, high intellect), and orderliness (high conscientiousness, low intellect). Although this may be a helpful approach to understanding personality, limitations of the model have been noted (Saucier & Ostendorf, 1999). The segments of the circle are divided into 90 segments of 30°, which assumes that the distribution of natural language terms is even, which empirically does not appear to be so. Furthermore, these terms within each set do not seem to be coherently related to each other. Despite this, it still seems to be an alternative approach to viewing personality as a hierarchical structural.

California Personality Inventory (CPI) Adopting items and empirical scale construction techniques from the original Minnesota Multiphasic Personality Inventory (Hathaway & McKinley, 1943), the CPI was developed to measure personality attributes relevant to everyday interpersonal behaviors of ordinary people. That is, it aims to measure "folk culture"—concepts people use to describe and understand one another throughout history and cross-culturally (Gough & Bradley, 1996). It contains 434 items capturing 20 folk culture scales, six of which potentially belong to the domain of conscientiousness, both conceptually and empirically: responsibility (adherence to rules), socialization (adherence to social

rules), self-control (emotion management and attention to detail), good impression (concern about external opinions), well-being (focus on health and future), and achievement via conformance (working within defined boundaries and expectations) (Roberts et al., 2005).

Hogan Personality Inventory (HPI) The HPI is informed by socioanalytic theory and primarily aims to predict job performance (Hogan & Holland, 2003). The current version comprises of seven primary scales, which conceptually align with the Big Five (Hogan & Hogan, 1992). Prudence is the primary scale that represents conscientiousness and consists of seven "homogenous item composites," conceptually similar to facets: (a) moralistic (adherence to conventions), (b) mastery (diligence), (c) virtuous (perfectionistic), (d) not autonomous (concern about external opinions), (e) not spontaneous (planned), (f) impulse control (self-control), and (g) avoiding trouble (obedient). Prudence, however, seems to be a somewhat different construct to conscientiousness, correlating at 0.36 (Pace & Brannick, 2010). In terms of its homogenous item composites, prudence seems to capture aspects such as conformity and concern for other's opinion of the self, which are more likely to form parts of the openness domain than the conscientiousness domain in the traditional Big Five conceptualizations of personality.

Jackson Personality Inventory—Revised (JPI-R) The JPI-R consists of 15 scales, which form into five higher-order clusters (Jackson, 1994). Dependable is the cluster most strongly resembling conscientiousness and consists of three scales: organization (planning and time management), traditional values (adherence to conventional values), and responsibility (moralistic).

Multidimensional Personality Questionnaire (MPQ) The MPQ includes 11 primary trait scales, which fall under three or four orthogonal higher-order factors (Tellegen, 1990). One of these factors is called constraint, which is similar to the Big Five conscientiousness. Like other conscientiousness facets, there are primary trait scales within the constraint factor, which are control (cautious), harm avoidance (avoids adventure and danger), and traditionalism (adherence to moral, religious, and societal norms and values).

HEXACO Personality Inventory (HEXACO-PI-R) The HEXACO includes four facets of conscientiousness: organization (organization of tasks and possessions), diligence (industriousness, a strong work ethic), perfectionism (detail-mindedness, carefulness), and prudence (deliberation and impulse control; Lee & Ashton, 2004). Content pertaining to the "*moral* conscience, such as *honest* or *sincere*" found in conscientiousness facets of other personality inventories is not included in the HEXACO's conceptualization of conscientiousness (Lee & Ashton, 2004, p. 337). These items would appear to be better matched to the honesty/humility domain within this model.

Conceptual Synthesis of Conscientiousness Facets Across Models Many of the abovementioned facets are the same or similar across multiple faceted models of conscientiousness. For example, there is a facet labeled "perfectionism" in the 16PF, HEXACO, and AB5C models and a facet called "organization" in the HEXACO, AB5C, HPI, and JPI models. Other comparisons across models show facets with different labels that are very conceptually similar (e.g., NEO-PI-R "self-discipline" versus CPI "self-control"). In Table 7.1, we have attempted a conceptual mapping of the different facets of these eight faceted models (and one aspect model) to illustrate where the points of consensus and dissension across models occur.

As Table 7.1 shows, dutifulness, order, and deliberation were the most commonly occurring facets—these were each included in five of the eight models, demonstrating some consensus that these form a core of conscientiousness. Note that these three facets collectively form the "orderliness" aspect, demonstrating that there is greater consensus among researchers on orderliness-related facets. While the three facets of the industriousness aspect were included in four of the eight different models of conscientiousness, the conceptual correspondence of facet labels and definitions was not as clear (e.g., while we included MPQ "harm avoidance" and HPI "mastery" as representative of a competence/self-efficacy factor, these clearly represent different interpretations, framing, or focus of competence). Perfectionism was included in three models, and factors representing a preoccupation with social norms ("good impression," "virtuous," "traditional values") were also included in four of the eight different models. Collectively, this conceptual mapping demonstrates both: (a) some degree of consistency across models, in terms of which specific facets underlie conscientiousness, and (b) that while some models are more comprehensive than others, there is not yet one single theoretical model that includes all eight of the potential facets that comprise conscientiousness.

7.2.3 Empirical Investigations of the Structure of Conscientiousness

While conceptual mapping such as we have undertaken in Table 7.1 can be a useful starting point for identifying consensus and dissension across multiple theories, a corresponding empirical synthesis is required as evidence for facet equivalence. To address this need for an empirical synthesis, various investigators have factor-analyzed multiple personality inventories with the aim of identifying the "true" structure of personality. Some of these results pertaining to conscientiousness can be found in Table 7.2.

As can be seen in Table 7.2, ten facets have been found across studies: orderliness, industriousness, responsibility, self-control, decisiveness, traditionality, persistence, punctuality, formality, and virtue. However, only four facets appear consistently across multiple investigations: *orderliness*, *industriousness*, *responsibility*, and *self-control*. The replicability of these facets indicates that a comprehen-

sive model of conscientiousness may be needed, which incorporates the four facets. Here, the facet of orderliness would refer to one's level of organization, tidiness, and meticulousness. Industriousness would capture one's level of purposefulness of behavior and ambition. Responsibility would assess one's level of reliability. This facet has often also contained aspects of punctuality (Saucier & Ostendorf, 1999). In fact, Roberts et al. (2004) questioned the status of his punctuality factor and regarded it as what Saucier (2002) would call a "parcel" factor—a small number of semantically similar items which aggregated as one factor. However, Jackson et al. (2010) also reported punctuality to be a separate facet, such that its status as a separate factor versus an underlying parcel of a larger responsibility facet is unclear. Lastly, self-control would measure one's level of care and cautiousness. Thus, if one aims to capture the essential elements of conscientiousness, their measurement of use needs to contain these four facets.

7.3 Related Constructs

There are several noncognitive constructs commonly used in education that were developed and conceptualized independently of the five-factor model of personality, but are conceptually and/or empirically similar to conscientiousness. These include grit, time management, perfectionism, motivation, self-regulation, goal theory (performance/mastery versus approach/avoidance goals), and mindset theory. Considering these constructs under the broad umbrella of conscientiousness allows links to be made between educational and personality psychology, facilitating cross-fertilization of research. That is, research on the facets of conscientiousness can inform education research. In the sections below, we focus on three constructs that have clear conceptual and empirical links to conscientiousness: (1) grit, (2) time management, and (3) perfectionism. We earlier alluded to the two key mechanisms underlying the conscientiousness/achievement link: motivational drivers and behavioral habits. Grit relates to the motivational elements (the internal drivers of achievement) whereas time management relates to the regular habits (the behavioral mechanisms that translate into achievement). Perfectionism may link to a possible "dark side" of conscientiousness that is rarely considered in prototypical conscientiousness research, but may be informative, particularly in an educational context.

7.3.1 Grit

Grit refers to one's perseverance and passion for long-term goals—gritty students will passionately persist in the pursuit of specific goals over months of years of sustained effort (Duckworth, Peterson, Matthews, & Kelly, 2007; Duckworth &

Table 7.2 Faceted structure of conscientiousness from seven studies

	Perugini and Gallucci (1997)	Saucier and Ostendorf (1999)	Peabody and De Raad (2002)	Roberts, Bogg, Walton, Chernyshenko, and Stark (2004)	Roberts et al. (2005)	MacCann, Duckworth, and Roberts (2009)	Jackson et al. (2010)
Model	NEO/BFQ	ABF5C adjectives	AB5C adjectives	AB5C adjectives	Seven (NEO-PI-R, 16PF, CPI, MPQ, JPI-R, HPI, ABC5C)	Twelve (16PF, AB5C, CPI, CHS Perfectionism Scale, HPI, HEXACO, JPI, MPQ, NEO-PI-R, 6FQ, TCI, VIA)	Behavioral Indicators of Conscientiousness (BIC)
Unit analyzed	Adjectives	Adjectives	Adjectives	Adjectives	Scale scores	IPIP items	Behavior words
Language	Italian	English, German	Six European languages	English	English	English	English
Sample	Undergraduates (40 %), employees (60 %)	US students, German community volunteers	Six studies: five countries in Europe	Mostly undergraduates (89 %)	Community sample	High school students	Undergraduates, community sample, graduate and undergraduate RAs
Analysis	EFA	EFA	Conceptual judgments	EFA	EFA	EFA/CFA	EFA
Facet							
Order (time/ tasks)	Meticulousness (+), Inaccuracy (-)	Orderliness	Orderliness	Punctuality	Order	Task planning	Organization, Punctuality
Order (possessions)				Orderliness		Tidiness	Cleanliness, Appearance

(continued)

Table 7.2 (continued)

	Perugini and Gallucci (1997)	Saucier and Ostendorf (1999)	Peabody and De Raad (2002)	Roberts, Bogg, Walton, Chernyshenko, and Stark (2004)	Roberts et al. (2005)	MacCann, Duckworth, and Roberts (2009)	Jackson et al. (2010)
Industriousness	Superficiality (–)	Industriousness	Work	Industriousness	Industriousness	Industriousness, Perfectionism	Industriousness, Laziness
Responsibility	Reliability (+)	Reliability	Responsibleness	Reliability	Responsibility	Procrastination, Refrainment	Responsibility, Avoid work, Antisocial
Self-control	Recklessness (–)		Impulse control	Impulse control	Self-control	Control, Cautiousness	Impulsivity
Decisiveness		Decisiveness-consistency		Decisiveness			
Traditionality				Conventionality	Traditionalism		
Persistence			Persistence			Perseverance	
Formality				Formalness			Formality
Virtue					Virtue		

Gross, 2014). That is, grit involves elements of perseverance, achievement drive, intrinsic motivation to achieve goals, and a long-term rather than short-term focus. Duckworth et al. (2007) found two major factors underlying grit: *consistency of interests* (maintaining a consistent set of goals, interests, and passions that do not change) and *perseverance of effort* (a persistent tendency to work hard toward goal completion). In a series of studies, grit predicted educational attainment, Ivy League GPA, lower attrition from military training, and success at the national spelling bee (Duckworth et al., 2007; Duckworth & Quinn, 2009). Grit has emerged as a popular new construct in education policy and practice, with substantial media attention.

However, grit is very strongly correlated with the existing noncognitive construct of conscientiousness (ranging from $r = 0.70$ to $r = 0.77$; Duckworth et al., 2007; Duckworth & Quinn, 2009). Despite this very strong relationship, Duckworth and Quinn (2009) argued that grit is a separate construct from conscientiousness, based on evidence that grit shows incremental prediction of educational outcomes beyond the effects of personality (Duckworth et al., 2007; Duckworth & Quinn, 2009). However, later evidence emerging from different research groups contradicted these findings for high school students—grit did *not* significantly predict any of several educational outcomes beyond the effects of personality (Ivcevic & Brackett, 2014; MacCann & Roberts, 2010). In fact, neither of these later studies found a significant relationship between grit and GPA even *before* controlling for other variables (though grit was significantly associated with other educational outcomes such as rule violations). Given the conceptual and empirical overlap between grit and conscientiousness, personality researchers are now suggesting that grit represents "a subcomponent of conscientiousness" (Roberts, Lejuez, Krueger, Richards, & Hill, 2014, p. 1321) or "a lower-level personality trait in the domain of Conscientiousness" (Ivcevic & Brackett, 2014, p. 29). That is, grit may be included as a facet of conscientiousness at the same level as facets such as "orderliness," "achievement striving," or "self-control" facets shown in Tables 7.1 and 7.2.

The primary distinction between the grit and other facets of conscientiousness concepts appears to be grit's focus on *long-term* goals, where effort and persistence are maintained over a period years. As such, empirical findings showing the incremental prediction of long-term outcomes but not high school grades may not be surprising. Grit appears to be a subset of conscientiousness pertaining to a long-term focus, conceptually similar to persistence or related facets. MacCann and Roberts (2010) found that both components of grit showed the overall strongest relationship to the perseverance facet of conscientiousness, supporting this idea. The conceptualization of grit as involving emotional investment in long-term goals suggests that the grit and conscientiousness research could be fruitfully integrated to consider the motivational and emotional processes that link conscientiousness to academic success.

7.3.2 Time Management

Time management is a set of acquired behaviors and associated attitudes (MacCann et al., 2012). It has been defined as "behaviors that aim at achieving an effective use of time" (Claessens, Van Eerde, Rutte, & Roe, 2007, p. 36). Theoretical models of time management include content such as setting goals, using time management aids (e.g., planners, e-reminders), being organized, having structure and routine, engaging in short- and long-term planning, awareness or attention to deadlines, and coping with temporal flow (e.g., Britton & Tesser, 1991; Liu, Rijmen, MacCann, & Roberts, 2009; Macan, 1994). Time management and conscientiousness are closely related, with correlations ranging from 0.57 to 0.77 (Liu et al., 2009; MacCann et al., 2012). Correlations at the facet level can be even higher. For example, organization facets of time management and conscientiousness inventories have correlated at 0.86 with each other (MacCann & Roberts, 2010).

However, factor analysis of the time use efficiency scale and conscientiousness items supports the empirical distinction between the two (Kelly & Johnson, 2005). Like grit, it is possible that time management may represent a facet of conscientiousness relating to specific organization of time and tasks. A factor representing time management emerged in three of the empirical investigations of conscientiousness facets (labeled "punctuality" or "task planning") as reported in Table 7.2. An alternative conceptualization is that time management represents the behavioral expression of the conscientiousness personality trait (especially its organizational aspect). That is, time management is one of the behavioral mechanisms by which conscientious tendencies translate into real-life achievement. MacCann et al. (2012) provide empirical evidence for this effect, showing that time management scores fully mediate the link between conscientiousness and academic achievement in a sample of part-time vocational education students.

7.3.3 Perfectionism

Perfectionism is often considered in terms of two higher-order factors: *perfectionistic strivings* (setting high standards) and *perfectionistic concerns* (being anxious and worried in one's thoughts and behavior; Stoeber & Otto, 2006). Perfectionistic strivings are associated with adaptive psychological profiles such as higher motivation and higher achievement whereas perfectionistic concerns are associated with maladaptive psychological profiles such as trait anxiety and depression (Dunkley, Blankstein, Zuroff, Lecce, & Hui, 2006; Suddarth & Slaney, 2001). In terms of personality traits, perfectionistic strivings are significantly associated with conscientiousness but unrelated to emotional stability whereas perfectionistic concerns show significant *negative* associations with both conscientiousness and emotional stability (Kim, Chen, Kleitman, MacCann, & Karlov, 2015). That is, much of the empirical link between perfectionism and conscientiousness concerns the adaptive rather than maladaptive elements of perfectionism. There is emerging evidence of a

third factor of conscientiousness representing orderliness (tendency to be systematic and neat) (Kim et al., 2015; Suddarth & Slaney, 2001). This new "order" element of perfectionism demonstrates the strongest association with conscientiousness ($r=0.61$) and can be differentiated from other adaptive elements of perfectionism identified by perfectionistic strivings (Kim et al., 2015).

Given the strong conceptual and empirical relationships between perfectionism and conscientiousness, there is some consensus that perfectionism may constitute a key element or underlying facet of conscientiousness. In fact, the *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, 2013) defines perfectionism as an aspect of extreme conscientiousness characterized by an insistence of flawlessness. Similarly, several personality models include perfectionism as a facet of conscientiousness (see Table 7.1), and one of the empirical studies also suggested that one of the eight facets of conscientiousness would be best labeled as perfectionism (MacCann et al., 2009; see Table 7.2). However, perfectionism clearly contains maladaptive elements that are normally considered separate from traditional conceptualizations of conscientiousness (Stoeber & Otto, 2006). Moreover, the scope of perfectionism often encompasses intrapersonal elements in addition to interpersonal elements. For example, the Frost Multidimensional Perfectionism Scale contains the factors of parental expectations and parent criticisms (Frost, Marten, Lahart, & Rosenblate, 1990), and the Hewitt Multidimensional Perfectionism Scale includes a socially prescribed perfectionism factor, representing individual differences in the belief that high standards are expected from others (Hewitt & Flett, 1991). When perfectionism was included as an underlying facet of conscientiousness, the maladaptive and intrapersonal elements were not present, indicating that the "conscientiousness facet" view of perfectionism is somewhat narrower than the wider perfectionism literature.

7.4 Conscientiousness and Academic Outcomes

As mentioned earlier in this chapter, conscientiousness is a personality domain most strongly associated with academic achievement, explaining between 5 % and 8 % of the differences in students' GPA (O'Connor & Paunonen, 2007; Poropat, 2009; Richardson et al., 2012). In fact, the effect size of this relationship rivals that of intelligence (Poropat, 2009). However, there are a few key issues that impact on the conscientious/achievement relationship:

(a) The granularity at which conscientiousness is considered (which can relate to the bandwidth-fidelity debate of reliability versus broad content coverage).
(b) The source of the conscientiousness ratings (students' self-ratings generally show smaller effects than observer ratings from teachers or parents).
(c) The target of the ratings (the majority of educational research focuses on student conscientiousness, but educator conscientiousness may also play a role in student achievement).

We discuss each of these issues below.

7.4.1 Bandwidth-Fidelity Debate

The bandwidth-fidelity debate, first discussed by Cronbach and Gleser (1957), captures the tension between assessing personality with great precision and reliability (at the specific, narrow level of personality facets) versus assessing personality to gain broader content coverage and maximum information (by assessing broad, overarching domains of personality). This debate began anew in organizational psychology in the 1990s (e.g., Ones & Viswesvaran, 1996), culminating in the recent meta-analysis that used the 6–2–1 model to predict different aspects of job performance (Judge et al., 2013). The prediction of job performance from conscientiousness can be summarized as follows. For all three types of job performance: (a) achievement striving was a stronger predictor than the other five facets; (b) the industriousness aspect was a stronger predictor than the order aspect; and (c) the broad domain was as strong or stronger a predictor than any single facet or aspect. There is as yet no equivalent meta-analysis to summarize these facet-level relationships in educational research. However, evidence indicates that the results are similar in terms of the importance of the achievement striving and industriousness elements of conscientiousness and that some facets are clearly more predictive of academic achievement than others.

First, facets representing industriousness predict outcomes more strongly than the broad domain. Compared to broad conscientiousness, the industriousness facet was a stronger predictor of both university grades (Rikoon et al., 2016) and high school absences (MacCann et al., 2009). Also, both HEXACO diligence and NEO-PI-R achievement striving predict university GPA more strongly than broad conscientiousness (although this was not the case for high school GPA; Noftle & Robins, 2007). Second, some facets of conscientiousness show stronger prediction than others. The industriousness facet shows a stronger association than other facets with a range of criteria (university GPA, teacher ratings of student's pro-social behaviors, class absences, and disciplinary infractions), while the tidiness facet generally shows the weakest association (MacCann et al., 2009; Rikoon et al., 2014). The NEO-PI-R facets of achievement striving, self-discipline, and competence show stronger prediction of academic achievement than other facets (Noftle & Robins, 2007; O'Connor & Paunonen, 2007). That is, the subset of facets that is most strongly predictive of academic outcomes is that belonging to the *industriousness* aspect of the 6–2–1 model, rather than the *order* aspect.

There is a further progression into the bandwidth debate, where a consideration of granularity is proposed for the academic criterion as well as the personality predictors. That is, broader academic criteria, such as GPA, can be decomposed into their elements, such as essay submissions, multiple-choice exam, class participation, and attendance on the basis that each of these distinct components may be a function of different facets of a factor (O'Connor & Paunonen, 2007). The extent to which different elements of conscientiousness predict achievement may depend on which elements of achievement are being considered. For example, Lievens, Buyse, and Sackett (2005) distinguished medical school achievement in science-based coursework from achievement in interpersonal elements of medical practice (i.e.,

they considered two parts of the criterion space for achievement in medical school). They found that a situational judgment test of interpersonal abilities predicted interpersonal elements but not the science-based coursework. Similarly, it is feasible that conscientiousness may be a stronger predictor of achievement elements involving sustained effort (e.g., mark on a major project) than of achievement elements primarily dependent on differences in cognitive ability or memory (e.g., performance on standardized tests).

7.4.2 Self-and Other Ratings of the Student

One factor that substantially moderates the correlation between conscientiousness and academic achievement is the source of the rating of conscientiousness. Conscientiousness items can be self-rated (e.g., "I work hard") or can be rated by observers (e.g., "This student works hard"; "My child works hard"; see Fig. 7.1 for an illustration of self-ratings versus other ratings). Comparisons between the two sources of ratings indicate that other ratings (i.e., ratings provided by parents, peers, or teachers) are a substantially stronger predictor of academic performance than self-ratings (Poropat, 2014a, 2014b). In primary education, the correlation rises from 0.28 (Poropat, 2009) when conscientiousness is self-rated to 0.50 when rated by others (Poropat, 2014b). Similar increases in correlation from self- to other ratings are observed for secondary school GPA (0.21–0.38) and university GPA (0.23–0.38) levels (Poropat, 2009, 2014b). The correlation between other reported conscientiousness and university GPA is particularly noteworthy—it is one of the only meta-analytic findings where the association with GPA is higher than that observed for intelligence (0.21; Richardson et al., 2012). Interestingly, there is no evidence that the type of observer (teachers, parents, or peers) affects the correlation with academic performance, which is somewhat surprising given the differential engagement of these raters with students' academic behaviors.

The reason for this incremental validity of other over self-ratings of conscientiousness is unclear, but a possible explanation arises from Vazire's (2010) model of asymmetries in knowledge of personality. Specifically, Vazire argued that self- and other raters will have similar observational access to behaviors, while self-raters will be more likely to vary their ratings to present themselves positively when the rating is socially desirable. Thus, self-raters may bias their ratings, making them less-valid predictors. A related argument follows from the focus of other raters upon external behaviors rather than thoughts and feelings, because behaviorally based criteria, such as academic performance (Campbell, 1999), should be better predicted by behaviorally based ratings of personality (Connelly & Hülsheger, 2012), because behaviors are the best predictors of behaviors (Ouellette & Wood, 1998). Recent research from MacCann, Lipnevich, Poropat, Wiemers, and Roberts (2015) supports the idea that self-ratings and other ratings have differential prediction based on the type of information used by the self and others to make personality ratings. MacCann et al. distinguished between conscientiousness facets with content related to *approaching goals* versus *avoiding errors*. For self-ratings, both types of facets were

Rating Scales: Self-Report, School Frame of Reference, and Parent-Report
(example items from MacCann et al., 2009)

Self-report (facet)	School Frame of Reference	Parent-Report
I work hard. (*industriousness*)	I work hard at school.	My child works hard.
I like to organize things. (*tidiness*)	I like to organize things at school.	My child likes to organize things.
I think before I speak. (*caution*)	I think before I speak at school.	My child thinks before s speak.
I like to plan ahead. (*task planning*)	I like to plan ahead at school.	My child likes to plan ahead.

Forced Choice
(example items from MacCann et al., 2009)

Consider the following four statements.

A. I work hard.
B. I like to organize things.
C. I think before I speak.
D. I like to plan ahead.

Which statement is *MOST like you*? ___
Which statement is *LEAST like you?* ___

Anchoring Vignette
(example from Mõttus et al. 2012a, p. 316: competence facet)

PART 1: SELF-RATING
How much are you
capable, efficient, competent 1 2 3 4 5 inept, unprepared?
PART 2: VIGNETTE RATING
Will often goes back on his promises or finishes his works in a hurry or completes things imperfectly. When someone allows Will to use his things they can be rather sure that Will will ruin the borrowed things or lose them completely.
How much is Will
capable, efficient, competent 1 2 3 4 5 inept, unprepared?

Situational Judgment Test (SJT)
(example item from Olaru, MacCann, Schneider, Wilhelm, & Roberts, 2015: dependability facet)

You have two projects: one for the Superblog Company and one for CerealNow. You find the Superblogproject very interesting, and you get along very well with the Superblog staff. The CerealNow project is much less interesting to you, and you find parts of it quite difficult. You do not know any of the CerealNow staff very well.

How likely are you to use each of the following options? (*1: Very unlikely to do this; 2: Somewhat unlikely to do this; 3: May or may not do this; 4:Somewhat likely to do this; 5:Very likely to do this*)

(A) Get the boring CerealNow assigned to one of your coworkers so you can focus on the Superblog project, which inspires you.
(B) Decide every morning which project you want to work on that day.
(C) Devote equal time and effort to both projects.
(D) Devote more time to the interesting Superblog project because you feel confident you can make something really great.
(E) Devote more time to the difficult and boring CerealNow project because it may take you longer to do a good job.

Fig. 7.1 Example items assessing conscientiousness illustrating six different item types

equally predictive of high school academic achievement. For other ratings (obtained from the students' parents), approach-related facets showed much stronger prediction of achievement. These results support the idea that observer reports show stronger prediction of outcomes and also support the idea that this may be due to the type of information available to observers. That is, approach-related facets may be more behaviorally based (and hence more observable) than error-avoidance facets (which may be more related to internal processes).

7.4.3 Teacher Conscientiousness

The history of interest in the personality profile of successful or effective teachers dates back from the 1940s (Dodge, 1943). Despite this large interest, limited attention has been placed on the personality of teachers under the five-factor model. Nevertheless, the limited existing research illustrates the importance of teacher personality in the classroom. Teacher conscientiousness has predicted an objective measure of academic achievement—secondary teachers' self-rated conscientiousness played the most important role in predicting student scores on the Texas Assessment of Knowledge and Skills, a statewide assessment of core content knowledge (Garcia, Kupczynski, & Holland, 2011). Furthermore, both self-ratings and student-ratings of teacher conscientiousness are associated with student evaluations of teaching effectiveness at university (Kim & MacCann, 2013). However, a study of elementary and high school teachers found that self-rated conscientiousness was *not* related to their teaching performance as measured by the National Board for Professional Teaching Standards assessment (Emmerich, Rock, & Trapani, 2004). Nevertheless, this study found two personality factors reliably predicted differences between teachers' areas of specialization. First, teachers with high levels of openness tended to specialize in humanities subjects. Second, a personality factor relating to firmness rather than softness in teachers' leadership style (which shared considerable conceptual overlap with conscientiousness) was higher among career/technical teachers than art teachers. Taken together, these three studies support the idea that teacher conscientiousness as well as student conscientiousness may impact on student educational outcomes. However, further investigations are required to properly assess the impact of teacher personality on a variety of educational outcomes in different levels of education.

7.5 Application of Conscientiousness in Education

7.5.1 Large-Scale Assessment and International Comparison

In 2012, measures of conscientiousness were included in the global assessment of educational progress and outcomes: the Programme for International Student Assessment (PISA). Thus, national comparisons of the extent to which conscientiousness is important to student competencies in mathematics, reading, and science can be obtained from the PISA data. For illustrative purposes, we will discuss two of the PISA scales: (1) a five-item rating scale assessing the perseverance component of conscientiousness (sample item: "I continue working on tasks until everything is perfect") and (2) a nine-item assessment of conscientiousness administered in a mathematics frame of reference (sample items: "I pay attention in mathematics class," "I keep my mathematics work well organized"). Using publically available data, we provide very simple analyses demonstrating that these scales relate to student behavior as well as student academic competencies.

Across the 300,000 pupils who completed such personality scales for PISA 2012, conscientiousness was related to mathematics outcomes, reading outcomes, and science outcomes. Specifically, the (mathematics-specific) conscientiousness scale correlated 0.18 with mathematics performance, and the perseverance scale correlated 0.11 with mathematics and 0.10 with both reading and science performance. Students who were regularly late for school had lower scores on conscientiousness and perseverance than students who were not, for which the effect size (Cohen's *d*) was 0.29 for perseverance and 0.58 for conscientiousness. Similar differences were obtained for students who regularly truanted from school compared to those who did not (0.24 for perseverance and 0.44 for conscientiousness).

Given that conscientiousness thus predicts achievement as well as some behavioral drivers of achievement, it may make sense for countries to consider and report student conscientiousness when considering explanations for achievement and to use these to inform educational policy. In fact, the United Kingdom is producing reports of this nature (Wheater, Ager, Burge, & Sizmur, 2014). However, one of the major issues for using conscientiousness assessments in multi-country comparisons is the potential for cultural differences in response styles to mask mean differences in scores. For example, some cultural groups may value modesty and humility more than others. This may affect the way that group members respond to rating-scale assessments. Methods such as anchoring vignettes have been developed to address these issues, and we discuss these in more detail in the next section.

7.5.2 Selection

Personality has been widely used for job selection for decades, with conscientiousness as the strongest personality predictor of job performance (e.g., Carless, 2007; Judge et al., 2013). In contrast, standardized tests of personality have rarely been used for selection into education, despite extensive research demonstrating that conscientiousness is a strong predictor of academic performance. This may be changing, however. Many universities use reference letters as part of the applications process and are starting to use standardized referee reports to supplement or replace the traditional reference letter format. These standardized methods are essentially observer reports of the applicant's personality. One such method is the "Personal Potentiality Index" (PPI; Kyllonen, 2008). Instead of writing a reference letter, potential referees rate the applicant on 24 items that assess six content domains that share a rough correspondence to known personality domains: (1) knowledge/creativity (similar to openness), (2) communication skills (similar to extraversion), (3) teamwork (similar to agreeableness), (4) resilience (similar to emotional stability/low neuroticism), (5) planning/organization (similar to conscientiousness), and (6) ethics/integrity (similar to honesty/humility and also shares conceptual overlap with some of the rule-adherence elements of conscientiousness). That is, these standardized recommendations are analogous to observer ratings of personality.

One potential advantage of using standardized ratings is that all referees must cover the same content domain, whereas reference letters may differ slightly from letter to letter. It can be unclear whether a content gap in a reference letter is an inadvertent oversight of the referee versus a deliberate omission that indicates poor candidate standing on the relevant content domain. A further issue with using personality assessments for selection is the potential for response distortion. While the use of observer reports rather than self-reports may ameliorate this problem to some extent, it is unlikely that this solves the problem altogether. It is not hard to imagine a referee that exaggerates the positive qualities of a likeable but scatterbrained student who is desperate to gain admission to a particular course or a secondary school teacher who is pressured to facilitate student admissions to prestigious institutions in order to boost the high school's reputation

7.6 Traditional and New Approaches to Assessing Conscientiousness

7.6.1 Self-Report Scales

The most common way to operationalize conscientiousness is with self-report rating scales, but there are several potential issues with this. First, accurate measurement requires psychological insight into one's own behavior and motivations. This can be a particularly important consideration for K-12 educational applications, where self-concept may still be developing in the earlier grades. In the very early grades, there may also be concerns about the required level of literacy needed for self-reports on text-based assessments and the appropriate age group to switch from the puppetized assessments used with very young children to youth versions of standard rating scales. Barbaranelli, Caprara, Rabasca, and Pastorelli (2003) found accurate assessment with 8-year-olds' self-reports on a Big Five assessment designed for children. Soto, John, Gosling, and Potter (2011) found that reliability was acceptable for 10-year-olds answering the Big Five Inventory online, but was lower than that of adults. Specifically, the mean reliability across the five personality domains was 0.75 for 10-year-olds (versus 0.83 for adults), and the mean reliability of personality facets was 0.54 (versus 0.67 for adults). However, this data was obtained with a general assessment developed for adults, rather than a personality test that was specifically designed for children.

Second, self-report assessments may be prone to response distortion, especially if used for high-stake selection. Evidence shows that people can and do fake high on conscientiousness tests. People can increase their conscientiousness scores by 0.89 standard deviations when instructed to fake good (Viswesvaran & Ones, 1999). Moreover, job applicants tend to score 0.45 standard deviations higher on conscientiousness than non-applicants, suggesting that people really do increase their scores when taking tests for selection purposes (Birkeland, Manson, Kisamore, Brannick,

& Smith, 2006). Item content that is obviously socially desirable is the most likely to be faked, and item content for conscientiousness is a clear target for faking. In fact, conscientiousness was the Big Five domain that showed the biggest mean difference in scores for job applicants compared to non-applicants, suggesting that this is the most commonly faked item content (Birkeland et al., 2006). If personality assessments are used for high-stake applications (such as selection into college, graduate school, or medical school), then self-report rating scales of conscientiousness may be inappropriate due to concerns about applicants faking high.

Third, the prototypical five-point rating scale with labels from "strongly agree" to "strongly disagree" can be interpreted quite differently by different people in different circumstances. Differences in scale interpretation can be broadly categorized into the test takers': (a) methods for "anchoring" their item ratings relative to some criteria and (b) response tendencies to answer a particular way, regardless of content.

The internal anchor used may be a particular *reference group* (e.g., rating "I work hard" by comparing myself to other students in the class) or a particular *frame of reference* (e.g., rating "I work hard" by considering my behavior at work). If there are genuine differences between groups, these can be obscured by a reference group effect. For example, if students rate themselves relative to others in a particularly conscientious class, then the mean score for the class on conscientiousness will be depressed. In this way, a conscientious class may score lower than a non-conscientious class on a conscientiousness rating scale. This reference group effect is of particular concern for group comparisons and can result in unusual effects at the aggregate level, as we outline in a later section.

Research on the frame-of-reference effect shows that specifying an education-related frame of reference for conscientiousness items increases reliability and the prediction of GPA (e.g., Lievens, De Corte, & Schollaert, 2008). Specifying a frame of reference may involve changing the instructions on a personality scale (e.g., "Rate these items in terms of your behavior *when you are at school*") or may involve adding a tag to each item (e.g., "*I work hard*" becomes "*I work hard at school*"; "*I try my best*" becomes "*I try my best at school*"). Lievens et al. (2008) argued that a frame-of-reference effect produces stronger relationships with criteria for two reasons. First, it reduces within-person inconsistency (i.e., all items are interpreted with the same frame of reference). Second, the predictor is more conceptually relevant (e.g., the test taker's conscientious behavior at school is a more relevant predictor of their school achievement than their conscientious behavior in general).

Response tendencies can include (a) an acquiescent response bias (agreeing with all items, regardless of content), (b) a defensive response bias (disagreeing with all items, regardless of content), (c) extreme responding (using the extreme scale points—strongly disagree and strongly agree—rather than the midpoints), and (d) neutral responding (using only the middle of the scale and avoiding the end points). There is evidence that different countries and cultures differ in these response tendencies and that this can obscure mean differences between groups (e.g., Chen, Lee, & Stevenson, 1995; Mõttus, Johnson, & Deary, 2012b; Van Herk, Poortinga, & Verhallen, 2004).

7.6.2 *Alternative Methods of Assessment*

Forced-Choice Approaches One of the methods for reducing response distortion is to present items in a forced-choice format. For example, a test taker might be asked, "Which is more like you: (a) I work hard or (b) I am kind to others?" (where option (a) represents conscientiousness and option (b) represents agreeableness). The test taker cannot fake high on both conscientiousness and agreeableness, as they must choose between two desirable alternatives. One of the long-standing problems with forced-choice assessments is that the scores produced are *ipsative*: scores on different attributes can only be compared within the same person, rather than between people, and so cannot be used for selection applications. Recent psychometric advances to scoring forced-choice assessments with item response theory (IRT) can produce non-ipsative scores from forced-choice data (Brown & Maydeu-Olivares, 2011, 2012, 2013). This important advance in scoring allows rating-scale tests to be administered in forced-choice formats for selection purposes (for a nontechnical demonstration of this method, see Anguiano-Carrasco, MacCann, Geiger, Seybert, & Roberts, 2015).

Anchoring Vignettes Anchoring vignettes are a method for separating individual differences in the use of the rating scale from individual differences in the construct of interest (e.g., conscientiousness). They were originally developed for applications in comparative health, political, and economic survey research (e.g., King & Wand, 2007; Salomon, Tandon, Murray, & World Health Survey Pilot Study Collaborating Group, 2004). Anchoring vignettes have recently been used for comparisons of personality across multiple countries, including large-scale educational assessments such as PISA (Kyllonen & Bertling, 2013) and comparisons of conscientiousness in adult samples (Mõttus, Allik, et al., 2012, Mõttus, Johnson, & Deary, 2012). Using this method, a test taker would first complete a rating-scale assessment in the standard way. Following this, they would read several vignettes describing hypothetical people and answer the same rating-scale questions about these hypothetical people. Differences in the vignette ratings represent differences in the use of the rating scale only, as the relevant content characteristics are encapsulated in the vignette and are the same across all test takers. These differences in rating-scale use can then be statistically partialled out of self-ratings, thus controlling for issues such as the reference group effect and response biases in rating-scale use. An example anchoring vignette is shown in Fig. 7.1.

In PISA 2012, anchoring vignettes were used to control for response style differences in ratings of teacher support in mathematics (see Kyllonen & Bertling, 2013). After adjusting for vignette ratings, the mean within-country correlation between teacher-support ratings and student achievement rose from $r = 0.03$ to $r = 0.13$. The change to between-country associations was even more dramatic. Before the anchoring vignette adjustment, the between-country association for teacher support and achievement was $r = -0.45$. That is, countries reporting greater levels of teacher support paradoxically showed lower achievement. After adjustment, the correlation

was reversed ($r = 0.29$), indicating that countries with more teacher support showed higher achievement (as might be expected). Anchoring vignettes have been used to adjust conscientiousness ratings when making national comparisons and show similarly dramatic changes to initially paradoxical results. For example, country-level conscientiousness ratings are negatively correlated with life expectancy and GDP, but these significant associations disappear after using anchoring vignettes to control for between-country differences in rating-scale use (Mõttus, Allik, et al., 2012).

Situational Judgment Tests (SJTs) In Situational Judgment Test (SJT) items, test takers are presented with a real-life situation and asked to evaluate several possible responses. SJT items may be presented as written text or in multimedia format and may use a variety of response formats (e.g., asking test takers to select the best response, to rate the effectiveness of each response, or to rate how likely they are to endorse each response). SJTs have been widely used to assess a variety of noncognitive constructs, primarily of interpersonal or socio-emotional skills. Evidence to date indicates that SJTs are a reliable and valid way to assess noncognitive competencies (e.g., Christian, Edwards, & Bradley, 2010). Recent research on SJTs suggests that this format is an accurate and useful way of assessing conscientiousness facets (Olaru, MacCann, Schneider, Wilhelm, & Roberts, 2015). Scenarios can be constructed such that possible responses vary in the extent to which they represent conscientious behavior (see Fig. 7.1 for an example). More conscientious test takers are more likely to endorse options representing high levels of conscientiousness. A 20-SJT item instrument assessing the dependability facet of conscientiousness has shown good evidence for construct validity (large correlations with three self-report assessments of conscientiousness, but only small- to trivial-sized correlations with other).

7.7 Developmental Trajectories for Conscientiousness

A variety of research programs have considered the developmental dynamics and underlying processes or genotype of personality, often referring to these more deep-seated aspects as temperament (McAdams & Olson, 2010; Rothbart, 2007; see Chap. 4). From this perspective, observed personality traits such as conscientiousness reflect the interaction of temperament factors with experience and environmental influences.

Most research on temperament has been with children and relied upon either observations or ratings of children by adults. For example, Thomas and Chess (1977; Thomas, Chess, & Birch, 1968), Thomas, Chess, and Birch (1968) developed a model of temperament by analyzing children's behaviors, while Buss and Plomin (1975) created a model using cross-species comparison but still relying on adult observations. More recently, Rothbart and her colleagues (e.g., Rothbart, 2007; Rothbart, Derryberry, & Posner, 1994) presented an integrated framework for temperament assessment and research, based on the idea that temperament reflects reactivity to stimuli along with effortful control processes adopted to moderate reactivity. Three broad dimensions dominate this framework: defensive reactions

resulting in negative affectivity such as fear and anger, approach reactions of activity and pleasure expressed in extraversion or surgency, and capacities for managing attention and inhibiting responses. This third dimension has been labeled effortful control, and Rothbart (2007) proposed that it was the temperament component that underpinned conscientiousness.

According to Rothbart (2007), effortful control is expressed through the control of attention and related forms of self-regulation. One expression of this is the inhibition of undesirable responses (Diamond, 2013), but effortful control is also expressed through maintaining focus on a specific plan of action while ignoring interfering distractions (Rothbart, 2007). Such control calls upon the executive attention network within the human brain (Posner & Rothbart, 2009), which appears to facilitate attention to low-intensity and delayed goals at the expense of short-term gratification or avoidance. It is these types of behaviors that have been linked with effortful control which in turn have led a series of reviewers to highlight the apparent similarities between effortful control conscientious behaviors. For example, with respect to learning and academic performance, some of the more important conscientiousness-related behaviors are those associated with attention to outcomes beyond those that are immediately available, such as by focusing on goals and plans and following social rules and norms (Roberts, Jackson, Fayard, Edmonds, & Meints, 2009). Consistent with this, De Pauw, Mervielde, and Van Leeuwen (2009) found that components of effortful control related to task persistence and attention control were particularly associated with conscientiousness among children. Consistent with this, Kanfer, Wolf, Kantrowitz, and Ackerman (2010) reported that variables related to effortful control such as self-regulation, motivation, and effort regulation form a consistent trait complex with conscientiousness.

7.8 Enhancing Conscientiousness

Although often assumed to be unchangeable, there is considerable evidence that not only can personality change (Roberts, Walton, & Viechtbauer, 2006), but that personality does change substantially in response to both general maturational processes and individual contexts and experience (Edmonds, Jackson, Fayard, & Roberts, 2008; Sutin, Costa, Wethington, & Eaton, 2010). Further, intentional practices intended to alter behavior have also been shown to lead to personality change (Roberts, 2006). Educators in particular have long sought to change enduring behavioral patterns and especially those associated with effort regulation (Graesser, 2009; Pashler et al., 2007). These endeavors are consistent with one of the central goals of education, the development of future employees for modern economies, requiring that students internalize work-relevant social norms such as rule following, dependability, and independent action guided by internalized values (Bowles, Gintis, & Meyer, 1999). Such efforts to make students more employable (Poropat, 2011) are at the same time likely to result in changes to their effort regulation and hence their rated levels of conscientiousness.

Attempts to improve skills and behaviors associated with conscientiousness have primarily been undertaken with preschool and primary school children. The Tools of the Mind program (Barnett et al., 2008; Diamond, Barnett, Thomas, & Munro, 2007) was produced from within a Vygotskian cognitive framework (Vygotsky, 1997). Although not entirely focused upon conscientiousness-related behaviors, children are taught skills of self-regulation and inhibitory control of impulses, habits, and distractions, which are directly related to Rothbart's (2007) account of effortful control and the specific components of effortful control that De Pauw et al. (2009) found to form a coherent factor with measures of conscientiousness. Enhancement of self-regulation in the Tools of the Mind program was associated with improved academic outcomes. Likewise, Tominey and McClelland (2011) found that a self-regulation intervention focused upon behavioral control had significant effects on self-regulation among children who were delayed in this area and improved academic performance for children generally. Given the evidence for long-term benefits resulting from enhancing what Heckman (2013) refers to as "non-cognitive skills," including the specific skills of self-regulation that help to underpin conscientiousness has considerable potential to enhance individual students' academic performance but also their broader life outcomes (Heckman, 2013).

A variety of approaches have been used for changing personality and associated factors that potentially may also ameliorate conscientiousness, either generically or by affecting specific facets. For example, attentional bias modification has demonstrable efficacy as a tool for modifying anxiety (Macleod & Mathews, 2012), in part because of its capacity for redirecting attention toward appetitive-motivational (Goetz, Robinson, & Meier, 2008) and goal-directed (Mischel & Ayduk, 2004) behavior. Given the goal-directed nature of much of conscientiousness, attentional bias modification seems likely to be useful for enhancing this trait (Cummings & Poropat, 2013). Somewhat comparable are techniques based upon evaluative conditioning (Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010), which has been shown to increase students' desire to engage in studying (Custers & Aarts, 2005). A very different possible approach relies upon mindfulness training, which has been linked with increases in attentional control (Chambers, Lo, & Allen, 2007) and self-regulation (Caldwell, Harrison, Adams, Quin, & Greeson, 2010). These findings are consistent with observed associations between mindfulness and conscientiousness (Siegling & Petrides, 2014) and also with individual achievement (Howell & Buro, 2010). So there are a variety of encouraging signs that conscientiousness, or at least conscientiousness-related variables, is modifiable in desirable directions for improving educational outcomes.

7.9 Conclusions

Since Webb's (1915) recognition of a factor later identified as conscientiousness, the evidence that this trait plays a central role in academic success has come to be reliably established and widely recognized. As outlined, practical application of the

knowledge about conscientiousness has been limited by deficiencies in both the understanding and measurement of this factor. These include ongoing debate regarding the facet structure of conscientiousness and its links with related traits, as well as problems with the mechanics of assessing any personality construct, especially those such as conscientiousness that are seen as socially desirable. A further issue is the extent to which conscientiousness is genetically based or developmentally malleable, which has implications for appropriate social and organizational responses. Thus, while current knowledge about the efficacy of conscientiousness with respect to life outcomes should encourage continued and growing attention, this should be informed by continued investigation of the underlying nature and measurement of conscientiousness.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*(5th ed.). Washington, D.C: American Psychiatric Publishing.

Anguiano-Carrasco, C., MacCann, C., Geiger, M., Seybert, J. M., & Roberts, R. D. (2015). Development of a forced-choice measure of typical-performance emotional intelligence. *Journal of Psychoeducational Assessment, 33*, 83–97.

Barbaranelli, C., Caprara, G. V., Rabasca, A., & Pastorelli, C. (2003). A questionnaire for measuring the Big Five in late childhood. *Personality and Individual Differences, 34*, 645–664.

Barnett, W. S., Jung, K., Yarosz, D. J., Thomas, J., Hornbeck, A., Stechuk, R., et al. (2008). Educational effects of the tools of the mind curriculum: A randomized trial. *Early Childhood Research Quarterly, 23*, 299–313.

Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment, 14*, 317–335.

Bowles, S., Gintis, H., & Meyer, P. (1999). The long shadow of work: Education, the family, and the reproduction of the social division of labor. *Critical Sociology, 25*, 286–305.

Britton, B. K., & Tesser, A. (1991). Effects of time-management practices on college grades. *Journal of Educational Psychology, 83*, 405–410.

Brown, A., & Maydeu-Olivares, A. (2011). Item response modeling of forced-choice questionnaires. *Educational and Psychological Measurement, 71*, 460–502.

Brown, A., & Maydeu-Olivares, A. (2012). Fitting a Thurstonian IRT model to forced-choice data using Mplus. *Behavior Research Methods, 44*, 1135–1147.

Brown, A., & Maydeu-Olivares, A. (2013). How IRT can solve problems of ipsative data in forced-choice questionnaires. *Psychological Methods, 18*, 36–52.

Buss, D. M., & Plomin, R. (1975). *A temperament theory of personality development*. New York: Wiley.

Caldwell, K., Harrison, M., Adams, M., Quin, R., & Greeson, J. (2010). Developing mindfulness in college students through movement-based courses: Effects on self-regulatory self-efficacy, mood, stress, and sleep quality. *Journal of American College Health, 58*, 433–442.

Campbell, J. P. (1999). The definition and measurement of performance in the new age. In D. R. Ilgen & E. D. Pulakos (Eds.), *The changing nature of performance: Implications for staffing, motivation, and development* (pp. 399–430). San Francisco: Jossey-Bass.

Carless, S. A. (2007). Graduate recruitment and selection in Australia. *International Journal of Selection and Assessment, 15*, 153–166.

Chambers, R., Lo, B. C. Y., & Allen, N. B. (2007). The impact of intensive mindfulness training on attentional control, cognitive style, and affect. *Cognitive Therapy and Research, 32*, 303–322.

Chamorro-Premuzic, T., & Furnham, A. (2005). *Personality and intellectual competence*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chen, C., Lee, S. Y., & Stevenson, H. W. (1995). Response style and cross-cultural comparisons of rating scales among East Asian and North American students. *Psychological Science, 6*, 170–175.

Chernyshenko, O. S., Stark, S., & Chan, K. Y. (2001). Investigating the hierarchical factor structure of the fifth edition of the 16PF: An application of the Schmid-Leiman orthogonalization procedure. *Educational and Psychological Measurement, 61*, 290–302.

Christian, M. S., Edwards, B. D., & Bradley, J. C. (2010). Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology, 63*, 83–117.

Claessens, B. J. C., Van Eerde, W., Rutte, C. G., & Roe, R. A. (2007). A review of the time management literature. *Personnel Review, 36*, 255–276.

Conn, S. R., & Rieke, M. L. (1994). *The 16PF fifth edition technical manual*. Champaign, IL: Institute for Personality and Ability Testing.

Connelly, B. S., & Hülsheger, U. R. (2012). A narrower scope or a clearer lens for personality? Examining sources of observers' advantages over self-reports for predicting performance. *Journal of Personality, 80*, 603–631.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEOI Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., McCrae, R. R., & Dye, D. A. (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO personality inventory. *Personality and Individual Differences, 12*, 887–898.

Cronbach, L. J., & Gleser, G. C. (1957). *Psychological tests and personnel decisions*. Urbana: University of Illinois Press.

Cummings, D., & Poropat, A. E. (2013). *Attentional bias modification for academic conscientiousness: A pilot study.* Paper presented at the Australian conference on personality & individual differences, Brisbane, QLD.

Custers, R., & Aarts, H. (2005). Positive affect as implicit motivator: On the nonconscious operation of behavioral goals. *Journal of Personality and Social Psychology, 89*, 129–142.

De Pauw, S. S. W., Mervielde, I., & Van Leeuwen, K. G. (2009). How are traits related to problem behavior in preschoolers? Similarities and contrasts between temperament and personality. *Journal of Abnormal Child Psychology, 37*, 309–325.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93*, 880–896.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135–168.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318*, 1387–1388.

Digman, J. M. (1989). Five robust trait dimensions: Development, stability, and utility. *Journal of Personality, 57*, 195–214.

Digman, J. M., & Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis, comparison, and interpretation of six major studies. *Multivariate Behavioral Research, 16*, 149–170.

Dodge, A. F. (1943). What are the personality traits of the successful teacher? *Journal of Applied Psychology, 27*, 325–337.

Duckworth, A.L., & Gross, J.J. (2014). Self-control and grit: Related but separable determinants of success. *Current Directions in Psychological Science, 23,* 319–325.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*, 1087–1101.

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the short grit scale (Grit–S). *Journal of Personality Assessment, 91*, 166–174.

Dunkley, D. M., Blankstein, K. R., Zuroff, D. C., Lecce, S., & Hui, D. (2006). Self-Critical and Personal Standards factors of perfectionism located within the five-factor model of personality. *Personality and Individual Differences, 40*, 409–420.

Edmonds, G. W., Jackson, J. J., Fayard, J. V., & Roberts, B. W. (2008). Is character fate, or is there hope to change my personality yet? *Social and Personality Psychology Compass, 2*, 399–413.

Emmerich, W., Rock, D. A., & Trapani, C. S. (2004). Multiple personality domains in relation to occupational choice and performance among established teachers. *ETS Research Report Series, 2004*(2), i-61.

Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *The Journal of Abnormal and Social Psychology, 44*, 329–344.

Frost, R. O., Marten, P., Lahart, C., & Rosenblate, R. (1990). The dimensions of perfectionism. *Cognitive Therapy and Research, 14*, 449–468.

Garcia, P. L. S., Kupczynski, L., & Holland, G. (2011). Impact of teacher personality styles on academic excellence of secondary students. *National Forum of Teacher Education Journal, 21*, 1–8.

Goetz, P. W., Robinson, M. D., & Meier, B. P. (2008). Attentional training of the appetitive motivation system: Effects on sensation seeking preferences and reward-based behavior. *Motivation and Emotion, 32*, 120–126.

Gough, H., & Bradley, P. (1996). *CPI Manual* (3rd ed.). Palo Alto, CA: Consulting Psychologists Press.

Graesser, A. C. (2009). Inaugural editorial for Journal of Educational Psychology. *Journal of Educational Psychology, 101*, 259–261.

Hathaway, S., & McKinley, J. (1943). *Manual for administering and scoring the MMPI*. Minneapolis, MN: National Computer Systems.

Heckman, J. J. (2013). *Giving kids a fair chance: (A strategy that works)*. Cambridge, MA: MIT Press.

Hewitt, P. L., & Flett, G. L. (1991). Perfectionism in the self and social contexts: Conceptualization, assessment, and association with psychopathology. *Journal of Personality and Social Psychology, 60*, 456–470.

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin, 136*, 390–421.

Hofstee, W. K., De Raad, B., & Goldberg, L. R. (1992). Integration of the Big Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology, 63*, 146–163.

Hogan, R. (1983). A socioanalytic theory of personality. In M. Page & R. Dienstbier (Eds.), *Nebraska symposium on motivation, 1982* (pp. 55–89). Lincoln, NE: University of Nebraska Press.

Hogan, R., & Hogan, J. (1992). *Hogan personality inventory manual*. Tulsa, OK: Hogan Assessment Systems.

Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology, 88*, 100–112.

Howell, A. J., & Buro, K. (2010). Relations among mindfulness, achievement-related self-regulation, and achievement emotions. *Journal of Happiness Studies, 12*, 1007–1022.

Ivcevic, Z., & Brackett, M. (2014). Predicting school success: Comparing conscientiousness, grit, and emotion regulation ability. *Journal of Research in Personality, 52*, 29–36.

Jackson, D. N. (1994). *The Jackson personality inventory-revised manual*. Port Huron, MI: Sigma Assessment Systems.

Jackson, J. J., Wood, D., Bogg, T., Walton, K. E., Harms, P. D., & Roberts, B. W. (2010). What do conscientious people do? Development and validation of the Behavioral Indicators of Conscientiousness (BIC). *Journal of Research in Personality, 44*, 501–511.

Jensen, A. R. (1980). *Bias in mental testing*. New York: The Free Press.

Judge, T. A., Rodell, J. B., Klinger, R. L., Simon, L. S., & Crawford, E. R. (2013). Hierarchical representations of the five-factor model of personality in predicting job performance: Integrating three organizing frameworks with two theoretical perspectives. *Journal of Applied Psychology, 98*, 875–925.

Kanfer, R., Wolf, M. B., Kantrowitz, T. M., & Ackerman, P. L. (2010). Ability and trait complex predictors of academic and job performance: A person-situation approach. *Applied Psychology: An International Review, 59*, 40–69.

Kelly, W. E., & Johnson, J. L. (2005). Time use efficiency and the five-factor model of personality. *Education, 125*, 511–515.

Kim, L. E., & MacCann, C. (2013). *The power of student perception: Student ratings of teacher personality predicts their satisfaction with the course and teacher.* Paper presented at the Australian conference on personality & individual differences, Brisbane, QLD.

Kim, L. E., Chen, L., Kleitman, S., MacCann, C., & Karlov, L. (2015). Evidence for three factors of perfectionism: Perfectionistic strivings, order, and perfectionistic concerns. *Personality and Individual Differences, 84*, 16–22.

King, G., & Wand, J. (2007). Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes. *Political Analysis, 15*, 46–66.

Kyllonen, P. C. (2008). *The research behind the ETS Personal Potential Index (PPI).* Background report. Princeton, NJ: Educational Testing Service.

Kyllonen, P. C., & Bertling, J. P. (2013). Innovative assessment methods to increase cross-country comparability. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (pp. 277–285). London: Chapman Hall/CRC Press.

Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO personality inventory. *Multivariate Behavioral Research, 39*, 329–358.

Lievens, F., Buyse, T., & Sackett, P. R. (2005). The operational validity of a video-based situational judgment test for medical college admissions: Illustrating the importance of matching predictor and criterion construct domains. *Journal of Applied Psychology, 90*, 442–452.

Lievens, F., De Corte, W., & Schollaert, E. (2008). A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology, 93*, 268–279.

Liu, O. L., Rijmen, F., MacCann, C., & Roberts, R. (2009). The assessment of time management in middle-school students. *Personality and Individual Differences, 47*, 174–179.

Macan, T. H. (1994). Time management: Test of a process model. *Journal of Applied Psychology, 79*, 381–391.

MacCann, C., Duckworth, A. L., & Roberts, R. D. (2009). Empirical identification of the major facets of conscientiousness. *Learning and Individual Differences, 19*, 451–458.

MacCann, C., Fogarty, G. J., & Roberts, R. D. (2012). Strategies for success in education: Time management is more important for part-time than full-time community college students. *Learning and Individual Differences, 22*, 618–623.

MacCann, C., Lipnevich, A. A., Poropat, A. E., Wiemers, M. J., & Roberts, R. D. (2015). Self- and parent-rated facets of conscientiousness predict academic outcomes: Parent-reports are more predictive, particularly for approach-oriented facets. *Learning and Individual Differences, 42*, 19–26.

MacCann, C., & Roberts, R. D. (2010). Do time management, grit, and self-control relate to academic achievement independently of conscientiousness? In R. Hicks (Ed.), *Personality and individual differences: Current directions* (pp. 79–90). Brisbane, QLD: Australian Academic Press.

Macleod, C., & Mathews, A. (2012). Cognitive bias modification approaches to anxiety. *Annual Review of Psychology, 8*, 189–217.

McAdams, D. P., & Olson, B. D. (2010). Personality development: Continuity and change over the life course. *Annual Review of Psychology, 61*, 517–542.

McCrae, R. R., & Costa, P. T., Jr. (1983). Joint factors in self-reports and ratings: Neuroticism, extraversion and openness to experience. *Personality and Individual Differences, 4*, 245–255.

Mischel, W., & Ayduk, O. (2004). Willpower in a cognitive-affective processing system. In F. Baumgarten & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 99–129). New York: The Guildford Press.

Mõttus, R., Allik, J., Realo, A., Rossier, J., Zecca, G., Ah-Kion, J., et al. (2012). The effect of response style on self-reported conscientiousness across 20 countries. *Personality and Social Psychology Bulletin, 38*, 1423–1436.

Mõttus, R., Johnson, W., & Deary, I. J. (2012). Personality traits in old age: Measurement and rank-order stability and some mean-level change. *Psychology and Aging, 27*, 243–249.
Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology, 93*, 116–130.
O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences, 43*, 971–990.
Olaru, G., MacCann, C., Schneider, R., Wilhelm, O., & Roberts, R. (2015). *In search of dependable situational judgment tests*. Paper presented at the International Society for the Study of Individual Differences, London, July 2015.
Ones, D. S., & Viswesvaran, C. (1996). Bandwidth-fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior, 17*, 609–626.
Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin, 124*, 54–74.
Pace, V. L., & Brannick, M. T. (2010). How similar are personality scales of the "same" construct? A meta-analytic investigation. *Personality and Individual Differences, 49*, 669–676.
Pashler, H., Bain, P., Bottge, B., Graesser, A. C., Koedinger, K., McDaniel, M., & Metcalf, J. (2007). *Organizing instruction and study to improve student learning*. Retrieved from http://ncer.ed.gov/
Peabody, D., & De Raad, B. (2002). The substantive nature of psycholexical personality factors: A comparison across languages. *Journal of Personality and Social Psychology, 83*, 983–997.
Peabody, D., & Goldberg, L. R. (1989). Some determinants of factor structures from personality-trait descriptors. *Journal of Personality and Social Psychology, 57*, 552–567.
Perugini, M., & Gallucci, M. (1997). A hierarchical faceted model of the Big Five. *European Journal of Personality, 11*, 279–301.
Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*, 322–338.
Poropat, A. E. (2011). The role of citizenship performance in academic achievement and graduate employability. *Education and Training, 53*, 499–514.
Poropat, A. E. (2014a). A meta-analysis of adult-rated child personality and academic performance in primary education. *British Journal of Educational Psychology, 84*, 239–252.
Poropat, A. E. (2014b). A meta-analysis of other-rated personality and academic performance. *Learning and Individual Differences, 34*, 24–32.
Posner, M. I., & Rothbart, M. K. (2009). Toward a physical basis of attention and self-regulation. *Physics of Life Reviews, 6*, 103–120.
Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*, 363–387.
Rikoon, S. H., Brenneman, M., Kim, L. E., Khorramdel, L., MacCann, C., Burrus, J., & Roberts, R. D. (2016). Facets of conscientiousness and their differential relationships with cognitive ability factors. *Journal of Research in Personality, 61*, 22–34.
Roberts, B. W. (2006). Personality development and organizational behavior. *Research in Organizational Behavior, 27*, 1–40.
Roberts, B. W., Bogg, T., Walton, K. E., Chernyshenko, O. S., & Stark, S. E. (2004). A lexical investigation of the lower-order structure of conscientiousness. *Journal of Research in Personality, 38*, 164–178.
Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology, 58*, 103–139.
Roberts, B. W., & Jackson, J. J. (2008). Sociogenomic personality psychology. *Journal of Personality, 76*, 1523–1544.
Roberts, B. W., Jackson, J. J., Fayard, J. V., Edmonds, G., & Meints, J. (2009). Conscientiousness. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 369–381). New York: Guilford.

Roberts, B. W., Lejuez, C., Krueger, R. F., Richards, J. M., & Hill, P. L. (2014). What is conscientiousness and how can it be assessed? *Developmental Psychology, 50*, 1315–1330.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Personality traits change in adulthood: Reply to Costa and McCrae (2006). *Psychological Bulletin, 132*, 29–32.

Rothbart, M. K. (2007). Temperament, development, and personality. *Current Directions in Psychological Science, 16*, 207–212.

Rothbart, M. K., Derryberry, D., & Posner, M. I. (1994). A psychobiological approach to the development of temperament. In J. E. Bates & T. D. Wachs (Eds.), *Temperament: Individual differences at the interface of biology and behavior* (pp. 83–116). Washington, DC: American Psychological Association.

Salomon, J. A., Tandon, A., Murray, C. J. L., & World Health Survey Pilot Study Collaborating Group. (2004). Comparability of self rated health: Cross sectional multi-country survey using anchoring vignettes. *British Medical Journal, 328*, 258–261.

Saucier, G. (2002). Orthogonal markers for orthogonal factors: The case of the Big Five. *Journal of Research in Personality, 36*, 1–31.

Saucier, G., & Ostendorf, F. (1999). Hierarchical subcomponents of the Big Five personality factors: A cross-language replication. *Journal of Personality and Social Psychology, 76*, 613–627.

Siegling, A. B., & Petrides, K. V. (2014). Measures of trait mindfulness: Convergent validity, shared dimensionality, and linkages to the five-factor model. *Frontiers in Psychology, 5*, 1164.

Smith, G. M. (1967). Usefulness of peer ratings of personality in educational research. *Educational and Psychological Measurement, 27*, 967–984.

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology, 100*, 330–348.

Stoeber, J., & Otto, K. (2006). Positive conceptions of perfectionism: Approaches, evidence, challenges. *Personality and Social Psychology Review, 10*, 295–319.

Suddarth, B. H., & Slaney, R. B. (2001). An investigation of the dimensions of perfectionism in college students. *Measurement and Evaluation in Counseling and Development, 34*, 157–165.

Sutin, A. R., Costa, P. T., Wethington, E., & Eaton, W. (2010). Turning points and lessons learned: Stressful life events and personality trait development across middle adulthood. *Psychology and Aging, 25*, 524–533.

Tellegen, A. (1990). *Manual for the multidimensional personality questionnaire*. Minneapolis, MN: University of Minnesota Press.

Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner-Mazel.

Thomas, A., Chess, S., & Birch, H. G. (1968). *Temperament and behavior disorders in children*. New York: New York University Press.

Tominey, S. L., & McClelland, M. M. (2011). Red light, purple light: Findings from a randomized trial using circle time games to improve behavioral self-regulation in preschool. *Early Education and Development, 22*, 489–519.

Van Herk, H., Poortinga, Y. H., & Verhallen, T. M. (2004). Response styles in rating scales evidence of method bias in data from six EU countries. *Journal of Cross-Cultural Psychology, 35*, 346–360.

Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology, 98*, 281–300.

Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement, 59*, 197–210.

Vygotsky, L. S. (1997). *The history of the development of higher mental functions*. New York: Plenum Press.

Wagerman, S. A., & Funder, D. C. (2007). Acquaintance reports of personality and academic achievement: A case for conscientiousness. *Journal of Research in Personality, 41*, 221–229.

Webb, E. (1915). *Character and intelligence: An attempt at an exact study of character*. Cambridge, MA: Cambridge University Press.

Wheater, R., Ager, R., Burge, B., & Sizmur, J. (2014). *Achievement of 15-year-olds in England: PISA 2012 national report (OECD programme for international student assessment)*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/299658/programme-for-international-student-assessment-pisa-2012-national-report-for-england.pdf

Wiggins, N., Blackburn, M., & Hackman, J. R. (1969). The prediction of first-year success in psychology: Peer ratings. *Journal of Educational Research, 63*, 81–85.

Chapter 8
Self-Concept: Determinants and Consequences of Academic Self-Concept in School Contexts

Ulrich Trautwein and Jens Möller

8.1 Introduction

"I'm pretty smart," "I'm really not a math person," "I'm good at learning foreign languages." Most readers will have an immediate reaction when they see these statements, their answers to each of the statements will range from complete disagreement to complete agreement, and most readers will not agree completely with all three statements.

The three statements are items from instruments that assess students' academic self-concept. Self-concepts are subjective beliefs about the qualities that characterize us, with academic self-concepts describing our self-beliefs about our intellectual strengths and weaknesses. Self-concepts do *not* correlate perfectly with objective measures of academic potential or achievement. It is this somewhat moderate association with achievement that makes self-concepts so important for educational practice and so interesting for researchers: Self-concepts are not an inner mirror of outside reality, but they still reflect a certain "reality": our own reality. And it is this personal, private reality—and not the objective reality—that is most closely related to what we think and consequently what we do. Feeling competent in a specific area motivates and energizes behavior in that domain and is associated with many

Lipnevich, A., Preckel, F., & Roberts, R. (Eds.), *Psychosocial skills and school systems in the twenty-first century: Theory, research, and applications*. Berlin: Springer.

U. Trautwein (✉)
Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Europastraße 6, Tübingen, Germany
e-mail: Ulrich.trautwein@uni-tuebingen.de

J. Möller
Department of Psychology, University of Kiel, Olshausenstraße 75, 24118 Kiel, Germany
e-mail: jmoeller@psychologie.uni-kiel.de

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_8

favorable long-term outcomes (Bandura, 1997; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005; Trautwein, Lüdtke, Kastens, & Köller, 2006).

In light of the association of self-concept with many favorable outcomes, it is not surprising that it is featured prominently in major theoretical accounts of human motivation, including expectancy-value theory (see Eccles, 1983), self-determination theory (Deci & Ryan, 2000), and Pintrich's (2003) delineation of motivational science (see also Chaps. 9 and 10). Moreover, there are several good books and hundreds of insightful articles that focus solely on self-concept. Of course, the present chapter will not aspire to providing anything close to a full account of what is known and not known about self-concept. Instead, we will attempt to answer some of the most pressing questions about the role of self-concept as a central construct in educational theory and practice: What is self-concept? What are the consequences of high or low self-concept? What are the determinants of high or low self-concept? What can be done to positively influence self-concept?

Several psychological disciplines have contributed to research on self-concept. The present chapter is primarily informed by research in educational psychology because of its specific relevance for educational practice. Furthermore, the dominant focus of the present chapter is on academic self-concept as one central domain of self-concept.

8.2 Self-Concept: Core Characteristics

8.2.1 Definition, Structure, and Measurement of Self-Concept

Self-concepts are individuals' mental self-representations. In educational psychology, the focus is on subjective beliefs about one's own strengths and weaknesses. People typically differentiate between their qualities in many different areas. For instance, a student may attest to having high ability in mathematics but lower ability in languages. For this reason, it becomes necessary to differentiate between several domain-specific self-concepts. Consequently, the majority of current conceptualizations of self-concept highlight its *multidimensional structure*. In what has become a very influential review of research on self-concept, Shavelson, Hubner, and Stanton (1976; see Fig. 8.1) differentiated between four large domains of self-concept: academic, social, physical self-concept, and emotional. Within the domain of academic self-concept, they further differentiated between self-concepts in various different subjects. Supporting this multidimensional conceptualization, Marsh and colleagues (for an overview, see Marsh & Craven, 1997) have accumulated evidence for the empirical separability of several facets of self-concept. For instance, the Academic Self-Description Questionnaire (ASDQ; Marsh, 1990) differentiates between a total of 14 domain-specific self-concepts plus a global academic self-concept.

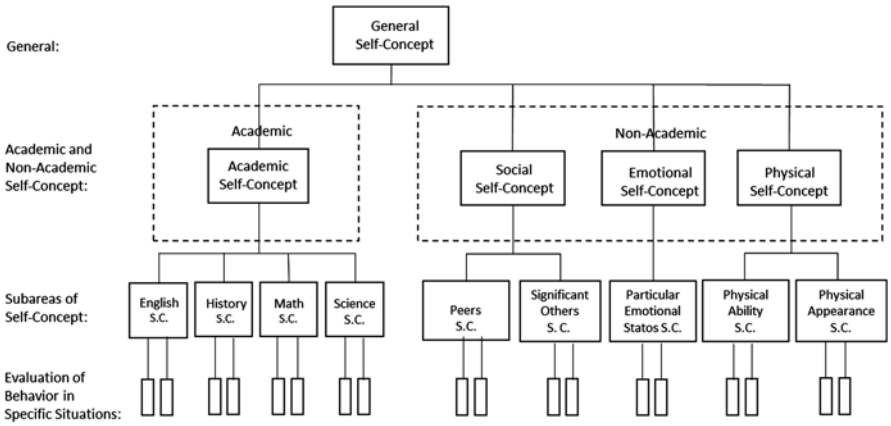

Fig. 8.1 The hierarchical self-concept model of Shavelson et al. (1976). *SC* Self-concept (Adapted from Shavelson et al. (1976, p. 413). Copyright © 1976 by SAGE journals. Reprinted by permission of SAGE Publications Inc)

In addition to the multidimensional nature of self-concept, the Shavelson et al. (1976) article also described self-concept as a hierarchy, in which a "general self-concept" formed the apex of the hierarchy. According to this conceptualization, self-concept at a higher (more general) level subsumes more domain-specific self-concepts. For instance, general academic self-concept was conceptualized as an integration of multiple domain-specific self-concepts. However, a number of empirical studies have found relatively low correlations between the self-concept facets, specifically for math vs. verbal self-concept. Accordingly, the self-concept hierarchy is, at best, a weak hierarchy. In fact, for academic self-concept, Marsh, Byrne, and Shavelson (1988) postulated an alternative conceptualization that differentiates between two largely independent second-order academic self-concept factors: verbal self-concept, which includes beliefs about academic abilities in first and second languages as well as history and geography and math self-concept, which integrates a student's beliefs about his or her abilities in subjects such as math, physics, chemistry, or biology. Figure 8.2 is a graphical depiction of this model. There is now substantial empirical support for this conceptualization of academic self-concept (see Marsh et al., 2015, for a recent application of this model). However, the academic debate about how to best describe the structure of academic self-concept has not yet reached its conclusion (e.g., see Brunner et al., 2010, for a nested factor model conceptualization of academic self-concept).

Self-concept is typically measured via self-report questionnaires (see Byrne, 1996, for an overview of instruments.) Students are presented with a list of statements (e.g., "I am good at math") and asked to indicate the degree to which they endorse the statements on a Likert-type scale. Typically, three items per domain are enough to achieve sufficient reliability and validity (Gogol et al., 2014). Whereas some of the early self-concept instruments conflated more affective ("I like

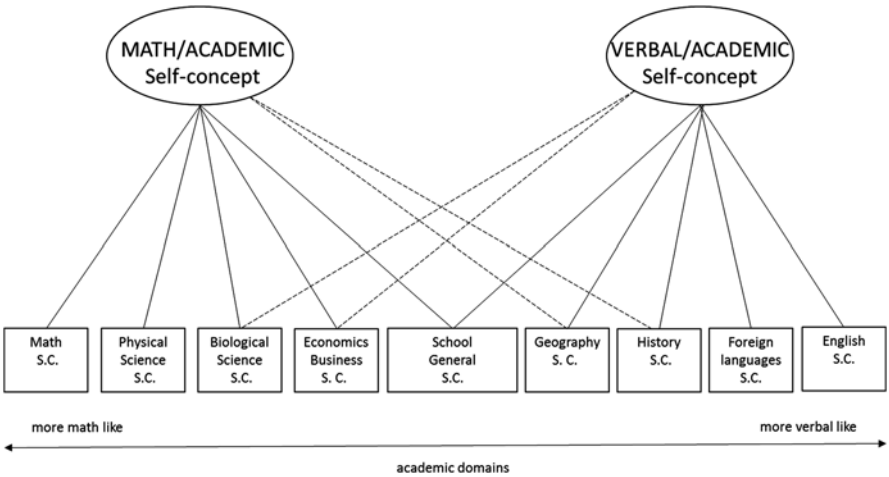

Fig. 8.2 The theoretical model of structure of academic self-concept developed as part of the Marsh/Shavelson revision (Marsh & Shavelson, 1985) of the Shavelson et al. academic self-concept model (1976) (Adapted from Marsh et al. (1988, p. 378). Copyright 1988 by the American Psychological Association. Adapted with permission)

English") and more cognitive-evaluative ("I have a talent for English") components of self-concept, the majority of modern approaches are restricted to the cognitive-evaluative aspect. Less agreement exists with regard to two other important conceptual differentiations of academic self-concept. First, some instruments seem to focus more on "accomplishments" ("I have always been good at school"), whereas other instruments highlight academic potential ("I have a talent for the sciences"). Second, the instruments differ in whether the items contain any comparison standards (see below) such as social ("I am one of the best students in my class") or dimensional ("English is one of my best subjects") comparisons. Despite these theoretically important differentiations, there seems to be a substantial empirical overlap between various self-concept instruments at the domain level.

8.2.2 Stability/Malleability of Self-Concept

From both theoretical and practical points of view, one important characteristic of self-concept is its assumed malleability: If self-concept is malleable, then educational interventions may be applied to foster more adaptive levels of self-concept. There are several different ways in which the stability of self-concept can be measured. Of specific importance are the stability of the level or degree of self-concept and the correlational (or normative) stability.

In terms of the *stability of the level of self-concept*, a number of studies (e.g., Harter, 1998; Marsh, 1989) that had used self-report questionnaires indicated that students' academic self-concept declines between school entry and adolescence. Whereas students in elementary classes often report a surprisingly positive evalua-

tion of their competencies, self-concepts become more realistic over time (Harter, 1998). Several factors—including cognitive development, feedback systems at school, and processes of developing one's own identity by differentiating domain-specific strengths and weaknesses—are likely to impact the process of self-concept development. By about age 15, mean self-concept scores seem to have stabilized at a much lower level than at school entry, followed by a small rebound to higher mean levels. The general decline in average self-concept scores does not imply that all students uniformly develop lower self-concepts in all school subjects. In fact, the general pattern of decline in self-concept somewhat masks a more differentiated pattern of self-concept development, in which many students have rather stable and high self-concepts in some subjects but markedly declining self-concepts in others.

Correlational (or the normative stability of) self-concept describes the *stability of interindividual differences in self-concept* across two or more measurement points. Domain-specific academic self-concept has been shown to be quite stable from its inception. For instance, Marsh, Craven, and Debus (1998) found 1-year stability coefficients of 0.46–0.64 for math, verbal, and academic self-concept in children 5–8 years of age. For older students, 1-year stabilities amount to 0.70 and higher (see Wigfield et al., 1997). In fact, the stability coefficients for academic self-concept are not very different from those for indicators of trait personality such as the Big Five (see Asendorpf & van Aken, 2003; see also Chap. 4). There are several plausible causes for this rather high correlational stability, including the stability of achievements, processes of selective attention to self-relevant information, and spirals of reinforcement (e.g., higher self-concept begets more effort and achievement, which in turn affect later academic self-concept). With regard to educational interventions, the rather high stability of academic self-concept under natural conditions can be considered a challenge.

8.3 Academic Self-Concept: Predicting Important Academic Outcomes

The interest in academic self-concept in research and practice is heavily rooted in the observation of existing links between academic self-concept and various key academic outcomes such as academic achievement, academic effort, and academic choices. This section briefly reviews the respective literature.

8.3.1 Academic Self-Concept, Academic Achievement, and Academic Choices

Academic self-concept and achievement are closely related. Domain-specific correlations between self-concept and school grades or test scores on standardized achievement tests typically amount to $r = \sim 0.40$ or higher (for a meta-analysis, see Möller, Pohlmann, Köller, & Marsh, 2009). Of course, one intriguing question in

self-concept research is about the causal direction: Does achievement determine self-concept (the so-called skill-development model; Calsyn & Kenny, 1977) or does self-concept affect later achievement (the so-called self-enhancement model; Calsyn & Kenny, 1977)?

Generally, in line with theoretical accounts of self-concept formation (e.g., Harter, 1998; Shavelson et al., 1976), there is agreement in self-concept research that academic achievement has an impact on the development of academic self-concept ("skill-development model"). There has been a more critical discussion with respect to the assumed effect of self-concept on achievement ("self-enhancement model"). A provocative and highly cited review by Baumeister, Campbell, Krueger, and Vohs (2003) seemed to suggest that self-concept has little or no effect on subsequent achievement. However, the Baumeister et al. article focused primarily on global self-esteem rather than on academic self-concept. The pattern of results reported in a number of other reviews or meta-analyses (Huang, 2011; Marsh & Craven, 2006; Marsh & Martin, 2011; Valentine, DuBois, & Cooper, 2004) has generally provided support for self-enhancement effects. Most notably, the meta-analysis by Valentine et al. (2004) provided a comprehensive overview of all studies ($N = 55$ publications, totaling more than 50,000 participants) that had examined the effects of prior self-beliefs on later achievement, controlling for prior achievement. They found an overall positive effect of self-concept on later achievement. Moreover, the effects were stronger when a study used an indicator of academic self-concept rather than global self-concept/self-esteem (for an example, see Trautwein, Lüdtke, Köller, & Baumert, 2006); similarly, the associations were stronger when the self-belief measure matched the achievement domain (e.g., math self-concept and math achievement).

Taken together, there is now robust empirical evidence from longitudinal studies indicating that both self-enhancement and skill-development effects take place. Consequently, the mutual effects that have been found for academic self-concept and achievement have been combined in the "reciprocal effects model" (Marsh & Craven, 2006; Marsh & Martin, 2011; Retelsdorf, Köller, & Möller, 2014).

Why is a comparatively high self-concept associated with a positive development of academic achievement? What are the underlying mechanisms? A number of studies have examined these processes and have found evidence for the mediating role of adaptive academic behavior (e.g., Helmke, 1990; Wigfield & Eccles, 1992). For instance, in Trautwein, Lüdtke, Marsh, and Nagy (2009, Study 2), competence beliefs strongly predicted academic effort, which, in turn, was associated with higher achievement on a posttest, even after controlling for prior achievement.

In addition to its positive effect on academic achievement, self-concept is also a good predictor of academic choices. For instance, higher self-concept in a specific domain predicts course choices, college major, and job choices (e.g., Eccles, 1994; Marsh & Yeung, 1997; Nagy, Trautwein, Baumert, Köller, & Garrett, 2006; Simpkins, Davis-Kean, & Eccles, 2006). Not surprisingly, the substantial effect of self-concept on academic choices has garnered considerable attention in many countries, in which there is a shortage of young adults who would like to pursue a career in STEM (science, technology, engineering, mathematics) fields (e.g., European Commission, 2007).

To sum up, academic self-concept is predicted by academic achievement and, in turn, predicts important outcome variables like motivation, course choices, and achievement, leading Marsh and Craven to the conclusion that it belongs to the most important constructs in the social sciences (Marsh & Craven, 2006).

8.3.2 Negative Consequences of an Overly High Self-Concept

Is a high academic self-concept always associated with adaptive outcomes? Doesn't an overly high self-concept rather reflect maladaptive cognitive or motivational strategies or problematic personality traits such as narcissism? In fact, critics (e.g., Dunlosky & Rawson, 2012; Kim, Chiu, & Zou, 2010; Kruger & Dunning, l999) have warned that an overly high academic self-concept may lead to overconfidence which in turn may be associated with underachievement. In this view, a well-calibrated self-concept is more adaptive. Empirical studies seem to support both positions: Negative effects of overconfidence have repeatedly been found in empirical studies, in which participants' judgments of their performance on a specific task were compared with their objective outcomes, whereas the findings from non-experimental longitudinal studies on the self (see Marsh & Craven, 2006; Valentine et al., 2004) have provided evidence that higher academic self-concept predicts a more positive development of academic achievement (and typically no support for negative quadratic effects). It is quite likely that the two research traditions are not measuring the same kind of self-evaluation: Academic self-concept is a rather stable and general self-evaluation of one's own abilities (e.g., "I am good at mathematics"), whereas measures in overconfidence research typically ask for evaluations of one's own performance on a specific task (e.g., "I'm sure I scored more points than anybody else today"). Nevertheless, there is the need for future research to develop a comprehensive model that can adequately explain the seemingly opposing empirical findings found for these different constructs.

8.4 Social, Dimensional, and Temporal Comparisons as Determinants of Academic Self-Concept

Self-concepts are determined by multiple factors. They do not directly reflect a student's "real" competence. For instance, a student who scores at the top of his or her class in English may still report that his or her English ability is nothing special. In other words, self-concepts do not directly correspond with outside criteria. The discrepancy between what is measured by outside criteria such as "objective" tests or observer reports and students' self-reports continues to trigger researchers' interest and research productivity in this field. In this section, we will describe social, dimensional, and temporal comparisons as three main determinants of academic self-concept.

8.4.1 Social Comparison Processes

A major determinant of self-concept is how we compare with the people around us: "Am I better or worse than my classmates, my peers, or my neighbors?" A day does not go by without such comparisons being made consciously or unconsciously, and psychological research on social comparison dates back to James (1892) and Festinger (1954). A large body of research suggests that people are constantly on the lookout for social comparison information that can be integrated into their self-concepts (Suls & Wheeler, 2000). Social comparisons occur when a person relates the abilities, opinions, or other characteristics of a person (which is often the person him- or herself) or group to the abilities, opinions, or other characteristics of another person or group. Why do people engage in social comparisons? Festinger (1954) suggested that evaluative information supplied by others serves the purposes of protection and survival of the self and hence provides a motivation for self-evaluation. More recent research has provided evidence that self-enhancement and self-maintenance (Tesser, 1988) also serve as motivations for comparisons with others.

As an additional self-improvement motivation, Festinger (1954) asserted that people preferred to compare their abilities to slightly better comparison standards ("There is a unidirectional drive upward in the case of abilities…" p. 124). In school, such upward social comparisons often seem to be caused by an aspiration to get hints from analyzing the results of students with better achievement. Conversely, there is also a strong support for the idea that upward social comparisons with better-performing others lead to more negative self-concepts (Wheeler & Suls, 2005, 2007). Downward social comparisons with worse-performing others often seem to be caused by the motivation to feel good or better; self-enhancement, self-maintenance, or self-protection motivations may trigger downward comparisons. Overall, the positive self-concept effect of downward comparisons is slightly higher than the negative self-concept effect of upward comparison (Möller & Pohlmann, 2010).

But what is the most important social comparison information in school, and what are the consequences of social comparisons? There is now ample evidence that a student's most immediate learning environment (i.e., his or her school class) is the most important frame of reference for the development of academic self-concept. The so-called big-fish-little-pond effect (Marsh, 1987) describes this phenomenon and its consequences.

8.4.1.1 The Big-Fish-Little-Pond Effect

One major strand of research in educational psychology has focused on the impact of achievement differences in naturally occurring educational environments (e.g., classes, schools) on outcome variables such as academic self-concept and educational choices (Marsh & Craven, 2002). In this paradigm, student outcomes are seen

as the consequence of specific characteristics of the (natural) learning environment. According to the big-fish-little-pond effect (BFLPE; Marsh, 1987), students typically compare their own achievement with the achievement of other students in their immediate learning environment. As a consequence, students with an average level of performance will develop a relatively high academic self-concept when they are placed in very low-achieving classes. They become a big fish in a small pond. By contrast, students with an identical level of performance will develop lower academic self-concepts when they are placed in high-achieving classes. This effect is partly mediated by the performance feedback provided by teachers. In high-achieving classes, students will get lower grades for the same objective achievement than students in low-achieving classes (Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006). Classes with high-achieving students offer more opportunities for upward social comparisons, which have negative consequences for individuals' academic self-concepts.

The analytical approach most frequently chosen in studies examining frame-of-reference effects in educational environments is regression based. General or domain-specific academic self-concept (assessed by items such as "I am smart" or "I am good at mathematics") is used as the outcome variable, and individual student achievement and school-average achievement are used as the two major predictor variables. Regression analysis is used to test whether school-average achievement is positively or negatively associated with self-concept when individual achievement is statistically controlled. In other words, it examines the consequences of placement in high- or low-achieving environments. Given two students with comparable achievement scores, which student has a higher academic self-concept: the one placed in a high-achieving school or the one placed in a low-achieving school?

The large majority of studies that employed regression analyses found a negative effect of school- or class-average achievement as measured by standardized achievement tests on academic self-concept (for reviews, see Marsh, Trautwein, Lüdtke, & Köller, 2008; Marsh & Craven, 2002). In fact, empirical support for the BFLPE is compelling. For instance, Marsh and Hau (2003) conducted a large cross-cultural test of frame-of-reference effects using data from the Program for International Student Assessment (PISA; OECD, 2001). Nationally representative samples of approximately 4,000 students from each of the 26 participating countries (total $N = 103{,}558$ students in 3,851 schools) completed standardized achievement tests and a self-concept questionnaire. Consistent with a priori predictions, the predictive effects of individual student achievement were substantial and positive, whereas the regression coefficients for school-average achievement were negative. In other words, if two students have comparable achievement, there is some likelihood that the one in the more selective academic environment will report a *lower* academic self-concept.

Students in Germany are assigned to different school tracks on the basis of their achievement at about age 10. The most common tracks are Hauptschule (low track), Realschule (middle track), and Gymnasium (high track). Some researchers have used track status rather than school-average achievement to predict self-concept. For instance, Schwarzer, Lange, and Jerusalem (1982) examined the effect of track

status on the development of academic self-concept after the transition to secondary school. The academic self-concepts of high-achieving students (who were placed in the high track) tended to decrease after the transition to secondary school, whereas the academic self-concepts of low-achieving students (who were placed in the low track) tended to increase, indicating that the negative effect of high-achieving classmates was stronger than any positive effect of high-track membership.

8.4.1.2 Counterbalancing Effects: Does the Ranking or Reputation of the School Predict Self-Concept?

The studies reported thus far indicate that self-concepts of students who are placed in academically selective schools or classes are negatively affected—a BFLPE or contrast effect. However, might self-perceptions also be enhanced by one's membership in high-achieving or positively valued groups? In the social psychology literature, there is sound evidence that people enjoy basking in the reflected glory of successful others (e.g., Cialdini & Richardson, 1980) and that self-perceptions may be enhanced by membership in groups that are positively valued by the individual (Diener & Fujita, 1997; Tesser, 1988). Adopting the term "reflected glory effects," (Marsh, 1987; Marsh, Kong, & Hau, 2000) argued that—theoretically speaking—students in academically selective schools might have more positive academic self-concepts by virtue of being affiliated with a highly selective educational program. In this sense, placement in a high-achievement group might be expected to positively affect students' global and domain-specific self-concepts by means of "assimilation effects" (see Marsh et al., 2000; Oakes, 1985; Seaton et al., 2008; also see Trautwein, Köller, Lüdtke, & Baumert, 2005). From a theoretical point of view, these reflected glory effects may weaken or fully counterbalance negative frame-of-reference effects.

All in all, the available research indicates that assimilative effects that counterbalance the contrast effects exist, but they are almost always smaller than the contrast effects (Marsh et al., 2008). Notably, a study by Marsh et al. (2000) indicated that students' academic self-concepts are not fully determined by their relative position in school, but, rather, their self-concepts also reflect their beliefs about the relative standing of their school. Similarly, a study by Trautwein et al. (2009) used both between-school and within-school approaches to investigate frame-of-reference and reflected glory effects in education, incorporating students' own perceptions of the standing of their school and class. Multilevel analyses with data from three large-scale assessments indicated that, given comparable individual achievement, placement in high-achieving learning groups was associated with comparatively low academic self-concepts. However, students' academic self-concept was not merely a reflection of their relative position within the class, but it was also substantively associated with their individual and shared perceptions of the standing of their class within their school.

Further evidence that students actively integrate various sources of information when using social comparison was found in a study by Chmielewski, Dumont, and

Trautwein (2013). These authors used data from PISA and demonstrated that although there was support for negative contrast effects in all school systems, the effects were weaker for students in course-by-course tracking rather than between-school and within-school streaming. This indicates that the reference group to which students compare themselves differs between types of tracking, with tracking status being more visible/more salient in course-by-course tracking.

With regard to appropriate instructional support for gifted learners, there is often a call for special tracks, pull-out programs, or enrichment programs. In this context, potential negative frame-of-reference effects on gifted students' self-concept as a consequence of high-achieving peer learners have triggered a number of studies. Dai and Rinn (2008) argued that findings from some studies on gifted student programs indicate that students in these selective academic programs may be less affected by negative frame-of-reference effects. Similarly, Makel, Lee, Olszewski-Kubilius, and Putallaz (2012) found no negative effects of a supplemental academic summer program for high-achieving students on these high achievers' academic self-concepts. In a German study that directly compared assimilation and contrast effects in a sample of 722 fifth grade students, Preckel and Brüll (2010) found the counterbalancing effects to be of similar size.

Taken together, much of the available evidence indicates that an individual's academic self-concept is typically negatively affected when she or he is placed in high-achieving learning environments, but this effect might be weaker (or nonexistent) under certain conditions (e.g., some forms of gifted education.)

8.4.1.3 Do Class Characteristics Interact with Individual Student Characteristics?

Another important issue in the study of reference group effects is whether the social comparison effects apply to all students in the same way. In other words, after controlling for individual ability, if there is a negative effect of school-average ability on student self-concept, is this effect the same for all students in a class? Are high- and low-achieving students in the same class equally affected by a high average ability of their reference group? Overall, a relatively small number of studies have investigated whether the BFLPE is similar at all ability levels, and such studies have yielded nonsignificant results or relatively small effects, the direction of which has not been consistent across studies. For instance, findings by Marsh, Chessor, Craven, and Roche (1995) and Marsh and Hau (2003) suggested that the BFLPE affects all levels of ability in a similar way. Marsh, Trautwein, Lüdtke, Baumert, and Köller (2007) tested interaction effects between school-average ability and individual ability in two samples of college-track (Gymnasium) high school students. Whereas there was no evidence for an interaction effect in the first sample, a negative interaction term in the second sample suggested that high-achieving students were more strongly affected by their placement in high-achieving schools. Conversely, in the above mentioned study by Trautwein et al. (2009) involving three different samples, the negative effects of being placed in high-achieving learn-

ing groups were weaker for high-achieving students in all three samples. In sum, although the results have been inconclusive to date, individual differences in ability represent a potentially important BFLPE moderator that warrants further consideration.

8.4.2 Dimensional Comparison Processes: The I/E-Model

Whereas the social comparison (Festinger, 1954) theory is well established, dimensional comparison theory (Möller & Marsh, 2013; Möller, Helm, Müller-Kalthoff, Nagy, & Marsh, 2015) is a relatively recent approach to describing the formation of individuals' self-concepts. The term *dimensional comparison*, first used by Möller and Köller (2001), makes reference to intraindividual comparisons of performance across domains (see Chiu, 2012). Dimensional comparison entails a single individual comparing his or her ability in a target domain with his or her ability in a standard domain (e.g., "How good am I in math compared with English?"). The most prominent effect of dimensional comparisons is that students with identical verbal (or math) ability differ in their verbal (math) self-concept when they differ in their math (verbal) achievement: Students who perform better in math than in verbal domains tend to have a lower verbal self-concept than do students with identical verbal ability but lower math ability. Vice versa, students who perform worse in math than in the verbal domain tend to have a higher verbal self-concept than do students with identical verbal ability but higher math ability. The effects of dimensional comparisons were first described by Marsh (1986) in his internal/external frame-of-reference model (I/E model, Fig. 8.3).

The I/E model describes the joint operation of dimensional and social comparison processes. Using an external frame of reference, students conduct social comparisons by comparing their achievement with the achievements of their classmates in order to construct academic self-concepts. For example, if students' verbal achievement is higher than that of classmates', their verbal self-concept is likely to be higher as well. Using an internal frame of reference, students conduct dimensional comparisons by comparing their achievement in a given subject with their achievements in other subjects. For example, in a situation in which students' verbal achievement is higher than their math achievement, verbal self-concept will benefit and his or her math self-concept will suffer from dimensional comparisons.

With regard to the I/E model, academic self-concepts are the result of social and dimensional comparisons. In a meta-analysis of 69 studies with a total of $N = 125{,}308$ students, Möller et al. (2009) revealed the findings presented in Fig. 8.3. The average correlation between math and verbal achievements was positive and strong ($r = 0.67$) and much higher than the average correlation between math and verbal self-concepts ($r = 0.10$). The four path coefficients in Fig. 8.3 are central to the evaluation of support for the I/E model. As a result of social comparison processes, the two horizontal paths relating math achievement to math self-concept (0.61) and relating verbal achievement to verbal self-concept (0.49) were substantial

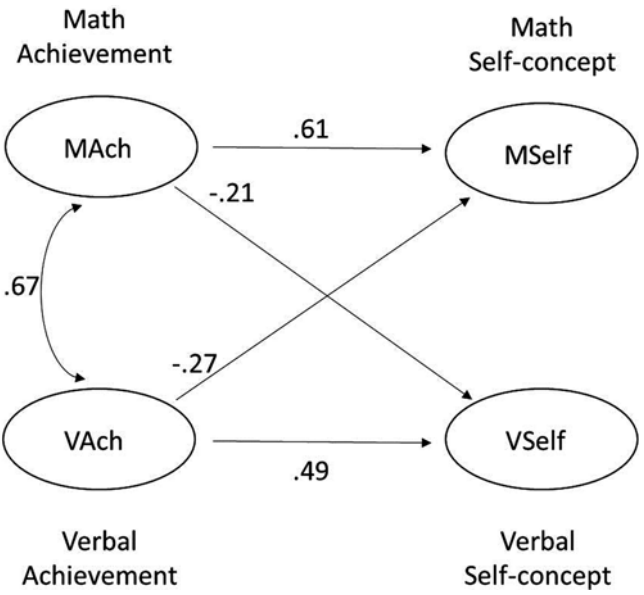

Fig. 8.3 The internal/external frame-of-reference model (I/E model): results of a meta-analytic path analysis on the relations between math and verbal achievement and math and verbal self-concept (Adapted from Möller et al. (2009, p. 1154). Copyright© 2009 by the American Educational Research Association. Reprinted by permission of SAGE Publications Inc)

and positive. As a result of dimensional comparison processes, the two cross paths leading from verbal achievement to math self-concept (−0.27) and mathematics achievement to verbal self-concept (−0.21) were negative. These findings indicate that the effects of dimensional comparisons described in the classic I/E model are not restricted to a particular achievement or self-concept measure or to specific age groups, gender groups, or countries (Möller et al., 2009; also see Pinxten et al., 2015).

Longitudinal studies measuring math and verbal achievement and math and verbal self-concepts on multiple occasions have further substantiated these findings (Marsh, Kong, & Hau, 2001; Marsh et al., 2014; Marsh & Yeung, 1998; Möller, Retelsdorf, Köller, & Marsh, 2011). Experimental studies have provided additional support for the idea that people evaluate their ability in one domain by comparing it with their ability in another domain (Möller & Köller, 2001; Möller & Savyon, 2003). Möller and Husemann (2006) conducted two introspective studies that confirmed that students spontaneously carry out dimensional comparisons in their everyday lives. Most of the dimensional comparisons employed in this study referred to academic matters, both as the target domain and as the standard domain ("We were given our school reports, and I compared my grade in religion with my grade in mathematics"). However, they also used a variety of other domains, specifically personal relationships and characteristics, as the standard domain ("This morning I thought: I am good at school, but when I compare that to making friends,

I seem worse"). Another important finding from this study, with regard to the motivation for dimensional comparisons, is that people seem to use dimensional comparisons to improve their self-worth and their mood.

8.4.3 Temporal Comparison Processes

Whereas social comparisons are based on interindividual frames of reference, temporal and dimensional comparison processes draw on intraindividual frames of reference. In 1977, Albert proposed temporal comparison theory—a complementary theory dealing with the temporal aspects of comparison processes. In temporal comparisons, a person compares his or her current performances and abilities with his or her own prior performances and abilities. Albert (1977, p. 488) stated that people try to maintain a coherent positive self-view. In their influential temporal self-appraisal theory, Ross and Wilson (2003) reviewed empirical evidence on temporal comparison and proposed that people tend to evaluate their past selves in a manner that makes them feel good about themselves now. According to Albert, people prefer to make comparisons with the near past, and temporal distance is negatively correlated with the tendency to carry out temporal comparisons. In Albert's view, temporal comparisons seem to have positive effects on self-concepts when comparisons with lower prior abilities are triggered but negative effects when comparisons with higher prior abilities are involved. Wilson and Ross (2001) added, however, that people can maintain their typically positive views of themselves under certain conditions by enhancing their former selves. To sum up, as with social and dimensional comparisons, temporal comparisons can have positive or negative effects on self-concept.

8.4.4 Multiple Frames of Comparison: A Complex Pattern of Influences

An exciting challenge for self-concept research is the fact that the various frames of reference are often in place simultaneously. For instance, imagine that Anna is a student who gets a B on her exam in English. This grade may carry social comparison information (the average grade of the class may have been between a B and a C) but also comparisons with a criterion ("an A is a good grade"). Furthermore, the "B" enables a temporal comparison (on her last exam, Anna received an A-) and a dimensional comparison (her last grade in mathematics was a C). The total effect of the exam on domain-specific English self-concept will depend on how Anna weighs all these pieces of comparative information and that is by no means easy to predict. Not surprisingly, the comparison processes (i.e., social, dimensional, temporal comparisons and comparison with a criterion) and their interdependence (see, e.g., Wakslak, Nussbaum, Liberman, & Trope, 2008; Wilson & Ross, 2000; Zell &

Alicke, 2009) continue to be the subject of empirical investigation and conceptual refinement.

Such a new conception is the reciprocal internal/external frame-of-reference model (RI/EM; Möller et al., 2011), which integrates social, dimensional, and temporal comparison information. It combines the I/E model with the reciprocal effects model by extending the I/E model longitudinally and the reciprocal effects model across domains. The model predicts that within domains, math and verbal achievement and academic self-concept will have positive effects on subsequent math and verbal achievement as a consequence of social and temporal comparisons. Across domains, math and verbal achievement and academic self-concepts should have negative effects on subsequent achievements and self-concepts as a consequence of dimensional comparisons (Möller, Zimmermann, & Köller, 2014; Niepel, Brunner, & Preckel, 2014).

8.5 Socializers' Feedback, Expectancy Effects, and Gender Stereotypes

Next to comparison processes, another major determinant of academic self-concept is direct and indirect feedback from parents and teachers as well as perceived societal norms. This is best exemplified by well-known gender differences in the development of self-concept. The academic self-concepts of boys and girls exhibit quite consistent differences that are fairly in line with prevailing gender stereotypes (Harter, 1998; Marsh & Hattie, 1996; Watt & Eccles, 2008). On average, boys report relatively higher mathematical self-concepts, whereas girls attest to having relatively more pronounced abilities in verbal domains. These gender differences reflect existing differences in achievement only in part.

Rather, such differences can be traced back to gender stereotypes that are expressed in the thoughts and actions of parents and teachers as the most important socializers. In a series of studies, Eccles and her colleagues (e.g., Frome & Eccles, 1998) were able to show that teachers and parents tend to believe that boys have more talent for math and science than girls—even if their actual performance is the same. Girls in turn tend to be perceived as higher on diligence and effort than boys (e.g., Tiedemann, 2000, 2002). Socializers' gender stereotypes were shown to impact students' domain-specific academic self-concepts, which in turn predicted academic effort, achievement, and academic choices (Frome & Eccles, 1998).

Gender stereotypes and their effects are perhaps the best-known examples of how teachers' or parents' expectancies can affect self-concept. However, there are similar findings in the literature on the Pygmalion effect (Rosenthal & Jacobson, 1968; see Jussim & Harber, 2005, for a review), indicating that a high evaluation of students' academic potential is associated with the specifically pronounced progress of these students. In this research, the most central dependent variable was usually achievement; however, effects on self-concept have also been studied and documented, and self-concept has been found to mediate the effects of expectancies on

achievement. From an educational perspective, an open question that still needs further research concerns the conceptual level at which expectancy effects take place (Friedrich et al., 2015): To what extent are expectancy effects taking place at the within- or between-classroom levels? If Pygmalion effects take place at the level of the whole class, it would be a potentially powerful instrument that could be used to enhance the self-concept and achievement of whole classes, whereas Pygmalion effects at the within-classroom level have some resemblance to zero-sum games (high progress in one student coupled with less progress in other students).

8.6 Self-Concept and Value Beliefs: Processes of Mutual Impact and Dedifferentiation

Students' interest in a specific domain has been identified as another potentially important determinant of academic self-concept. The underlying rationale is as follows: If a student is interested in a subject, he/she may engage in a specific way with the subject and develop higher skills, which, in turn, will have a positive effect on this student's self-concept in the specific domain. However, one could also argue for the opposite causal direction, in which students with high self-concepts in one domain also develop higher interest in this area (see Deci & Ryan, 2000; Eccles, 1994; Wigfield & Eccles, 1992). The question of "what comes first" or, more broadly, "what is the mutual impact of these constructs" is quite relevant for the educational practice, for instance, in relation to interventions. In this section, we describe the association between self-concept and interest, their mutual impact on each other, and their unique and shared effects on academic outcomes. To this end, we build on the Eccles (1983) expectancy-value model (see also Chap. 10).

8.6.1 Expectancy-Value Theory

According to the Eccles et al. (1983) expectancy-value model and its refinements (see Eccles & Wigfield, 2002; Wigfield & Eccles, 2000, 2002), expectancies and values directly influence performance, persistence, and task choice. In this model, the *expectancy of success* is conceptualized as a task-specific belief about success in a future academic task. Eccles et al. (1983) defined beliefs about ability as individuals' evaluations of their competence in different areas. Conceptually, the model differentiates between academic self-concept—i.e., broad beliefs about competence in a given domain—and expectancies of success on a specific upcoming task. However, in empirical studies, the two components have shown very high intercorrelations, and self-concept and expectancy beliefs have typically been collapsed into a single construct or used interchangeably (see Eccles & Wigfield, 2002). Another core aspect of the model is the differentiation of four task value components: intrinsic (or interest), attainment, utility, and cost (Eccles & Wigfield, 2002). *Intrinsic value* is defined as the

enjoyment a person derives from performing an activity or his/her subjective interest in a subject. *Attainment value* is defined as the personal importance of succeeding on a task. *Utility value* indicates the perceived (future) individual usefulness of engagement and achievement in a certain domain. Finally, *cost value* describes the perceived negative consequences of engaging in a task (e.g., performance anxiety, fear of failure, effort required, and the opportunity cost of choosing that option). Although theoretically separable, the four value components have shown relatively high intercorrelations in several empirical studies and have thus often been incorporated into a single, more general value scale (e.g., Eccles et al., 1983; see also Chap. 10).

8.6.2 Associations Between Self-Concept and Value Components

Both self-concept and value beliefs are highly domain specific (Bong, 2001; Eccles et al., 1993). Students often have favorite school subjects and will happily volunteer reasons for preferring mathematics over sports even if their objective performance in the two subjects is the same. In fact, correlations between self-concept and value beliefs across different school subjects are typically much lower than are correlations between the corresponding grades or test scores (e.g., Denissen, Zarrett, & Eccles, 2007; Möller & Marsh, 2013). Empirical evidence indicates that domain specificity of self-concept and value beliefs increases with students' age (Denissen et al., 2007). At the same time, the associations between expectancy and value beliefs within a domain increase over time (Denissen et al., 2007; Wigfield et al., 1997) because "children come to value what they are good at" (Wigfield et al., 2009; also see Archambault, Eccles, & Vida, 2010). Using a sample of students at the end of their school careers, Trautwein et al. (2012) reported associations between self-concept and intrinsic value as high as 0.80 for mathematics. In fact, the correlation between self-concept and intrinsic value was higher than some of the intercorrelations between value domains. This supports the notion of a particularly strong tie between what we believe we *can* do (self-concept) and what we *like* to do (interest).

What comes first? In general, there is some evidence for a mutual, reciprocal influence of domain-specific self-concepts and interests, even though more empirical studies are needed to see if there the strength of each effect is similar across different age groups and domains that are being studied (e.g., Marsh et al., 2005; Pinxten et al., 2015).

8.6.3 Predicting Academic Outcomes

The close association between self-concept and interest raises the question of whether researchers can possibly retain their focus on one and not attend to the other. Apart from conceptual differences between the two constructs, there are also empirical studies that have highlighted their unique effects, albeit not in all available studies. On a correlational basis, both expectancy and value beliefs tend to be

substantially associated with various academic outcomes. When both predictors are simultaneously entered into a prediction model, however, a different picture emerges. Whereas self-concept has been shown to be closely associated with performance in both cross-sectional and longitudinal studies, value beliefs tend to be more potent predictors of choice, effort, and persistence in achievement-related activities (see Wigfield & Eccles, 2000, 2002), although the pattern of results is not always clear-cut. With regard to students' homework behavior, (Trautwein, Lüdtke, Kastens & Köller, 2006; Trautwein & Lüdtke, 2007) found that intrinsic value was closely associated with the effort put toward homework; interestingly, a measure of self-concept/expectancy negatively predicted time spent on homework which could indicate that high self-concept is negatively associated with a dillydally working style. Conversely, in Meece, Wigfield, and Eccles (1990), value beliefs did not predict achievement after controlling for expectancy beliefs, although they were positively related to achievement in a simple univariate analysis. With regard to educational choices such as college major or course choices, value beliefs seem to be a better predictor, at least in modern Western societies (Watt & Eccles, 2008), but academic self-concept also has predictive power.

Moreover, as indicated by a number of recent articles, academic self-concept and interest may positively interact with each other in predicting academic outcomes. In other words, self-concept and interest do not just add to each other, but rather show a synergistic pattern: at high levels of both self-concept and interest, the predicted value for the outcome variable is especially pronounced (see Trautwein et al., 2012). For instance, Nagengast et al. (2011) tested the so-called expectancy-value interaction with PISA 2006 data, predicting science outcomes for large representative samples of 15-year-olds ($N = 398{,}750$; 57 countries). Expectancy (science self-concept), value (enjoyment of science), and the expectancy × value interaction all had statistically significant positive effects on both engagement in science activities and intentions to pursue scientific careers. These findings were similar for the total sample and for nearly all of the 57 countries.

8.7 Educational Implications

8.7.1 High Academic Self-Concept: A Powerful Resource

The previous sections offer a number of general implications. First and foremost, a high academic self-concept is a powerful resource for students. As described in Chap. 10 (this volume), academic self-concept is one of several competence-related constructs such as self-efficacy, expectancies, perceived control, and attributions. Compared with these other constructs, what is characteristic of self-concept is its multidimensional nature, the many frames of reference and comparison processes that impact its development, the mutual associations with interest development, the proximity of self-concept differentiation to identity ("What are my real strengths?"), and its predictive power for long-term academic choices. Accordingly, there is a

consensus that in general, (a) positive academic self-concepts protect students from losing interest in school and from becoming involved in problematic peer contexts and behaviors and (b) it is part of an educator's job to foster a (relatively) high academic self-concept in students.

Second, somewhat in contrast with psychosocial skills (e.g., self-regulation) that can and should be actively fostered in a step-by-step approach by teachers or specific training programs, self-concept development "happens" automatically all the time (see Chaps. 9, 12, and 13, for comparison). For instance, classrooms have been called a "total environment" that constantly provides social comparison information that is difficult to filter out. In addition, didactical and, perhaps, ethical questions ("Which academic self-concept domain should I foster?") may come into play when educators think about "appropriate" levels of academic self-concept. Hence, educational reforms or educational interventions that aim to foster students' self-concepts deal with a complex pattern of issues that need to be taken into account, including social and dimensional comparisons. In fact, reforms or interventions that improve some students' academic self-concepts may come at the price of having negative side effects on others.

Third, when it comes to the question of how to improve students' academic self-concept, there are two general routes. One, instead of targeting self-concept per se or suggesting to students that they are "full-grown geniuses," interventions should target the belief that one is "talented" in a certain domain or at school in general and simultaneously emphasize the need to develop the talent further. Two, there is reason to believe that academic self-concept can be improved indirectly by interventions that have different targets (e.g., self-efficacy, value beliefs, attributions)—perhaps as effectively or more effectively than via interventions that target self-concept directly (but see O'Mara, Marsh, Craven, & Debus, 2006).

Echoing and providing some examples for these more general conclusions, in the following section, we present a short and by no means comprehensive overview of empirical evidence and educational implications with regard to educational structures, teaching quality, and psychosocial interventions.

8.7.2 Changing Educational Structures

The major argument for or against a differentiated school system with several more or less explicit tracks is the presumed effect of such tracking on achievement. However, potential effects on academic self-concept have also always been an issue in the detracking movement in many school systems across the world: One of the hopes was that detracking would have a positive impact on low-achieving students' academic self-concepts (e.g., Oakes, 1985). Unfortunately, however, as the results of many studies on the big-fish-little-pond effect (Marsh, 1987; see above) have suggested, less tracking may not bring with it only positive effects for lower-achieving students but could in fact have negative side effects on their academic self-concepts (and, for that matter, positive effects on the academic self-concepts of

high-achieving students). Less tracking means greater heterogeneity in the student body, thus leading to lower-achieving students being confronted with higher-achieving students. In other words, social comparisons in detracked systems lead to less positive feedback for low-achieving students and more positive feedback for high-achieving students. Hence, it is therefore especially important for more comprehensive school systems that schools and teachers use instructional practices and provide psycho-social niches that bolster the academic self-concepts of low-achieving students.

Another caveat concerns the organization of secondary schools and the emphasis that is placed on competition. It has been suggested that the decline in student motivation from elementary to secondary schooling commonly found in national and cross-national studies is actually related to systematic changes in the classroom environment. It is assumed that the classroom environment in secondary schools fails to meet the specific developmental needs of students (Eccles et al., 1993). Students' increase in self-awareness and their increased need for independence in adolescence appear to be mismatched with the organization of the classroom environment, which is shaped by a high level of control and the promotion of social comparison and competition in secondary schools. In a somewhat similar vein, the emphasis on social comparison and competition that has been associated with an increase in the use of high-stakes testing has also become a concern for motivational researchers. For instance, Ryan and Weinstein (2009) argued that high-stakes testing may increase the use of controlling motivational strategies that undermine intrinsic motivation, and high-stakes tests (which provide something like an objective "criterion") may in fact undermine some students' self-concepts if they are experienced as too challenging.

8.7.3 Appropriate Cultures of Feedback: The Impact of Teachers

In the majority of empirical studies that study between-class differences in academic self-concept, the proportion of variance in academic self-concept that is accounted for by the classroom (the intraclass correlation coefficient) is rather moderate (e.g., Trautwein, Lüdtke, Kastens & Köller, 2006), indicating similar overall means of academic self-concept across classes. These findings support the notion that social comparison processes are constantly taking place in classes, making self-concept development—in the extreme expression of this—a zero-sum game. Research has yet to determine the extent to which the mechanisms that lead to such a zero-sum game can be overcome, but a number of possible solutions have already found some empirical support.

First, it has been argued that teachers should use temporal comparisons more often (see above) rather than social comparisons when giving achievement feedback. In other words, teachers should emphasize intraindividual, temporal comparisons that would help students appreciate the progress they have made. Such comparison

orientation has been shown to be associated with a more positive development of academic self-concept, even if it might not be able to eliminate social comparison processes (e.g., Lüdtke, Köller, Marsh, & Trautwein, 2005).

Second, social comparison processes with high-achieving students may be experienced as less damaging for self-concept if it is primarily constructed as a positive motivational cue ("This achievement level is possible if you invest the appropriate effort") or accompanied by informational value ("This student used a specific learning strategy that might be useful for everybody").

Third, a somewhat related issue is teachers' guidance in the use of supportive attributions for success and failure. Successes and failures may have a strong effect on academic self-concept when they are attributed to being caused by internal-stable factors such as (a lack of) talent. Especially in the case of failure, the use of internal variable attributions about the cause (e.g., lack of effort) is more beneficial because these causes are variable and controllable and the respective attribution promises improvement; for teachers, however, it is sometimes challenging to evaluate the appropriateness of such an attribution.

Fourth, in recent research on teaching effectiveness (e.g., Fauth, Decristan, Rieser, Klieme, & Büttner, 2014; Pianta & Hamre, 2009), the type of relationship that teachers foster with the students in their classroom has been shown to be connected to the development of student motivation. The provision of an emotionally safe and trusting environment, in which teachers act as role models as well as confidants who provide social support, has been found to have a positive impact on various dimensions of motivational predispositions, including self-concept. Moreover, as emphasized above, the development of reasonably high academic self-concepts is often the positive side effect of high-quality teaching more generally.

8.7.4 Psychosocial Intervention Studies

This final section addresses intervention research in regular classrooms or other academic environments. Such interventions more or less explicitly focus on the improvement of academic self-concept and can be implemented by teachers or administered by researchers. In the assessment of such measures, however, as described above, it is important to always keep in mind that self-concept development is affected by several frames of reference. Social and dimensional comparisons necessarily limit the strength and sustainability of interventions that target self-concept.

Despite these caveats, self-concept interventions may be quite successful. O'Mara et al. (2006) conducted a meta-analysis of 145 studies to describe the effects of intervention programs on the self-concepts of children and adolescents. In this meta-analysis, both academic and nonacademic self-concepts were included; furthermore, the meta-analysis included both domain-specific and domain-general measures of self-concept. Overall, an average effect size of $d = 0.47$ was found.

Accordingly, the training groups, in which the intervention was performed, had self-concepts that were nearly half a standard deviation higher than the control group. Interventions that were focused on a specific aspect of self-concept were particularly effective when this specific aspect of self-concept was also measured.

Despite these encouraging results, we argue that a crucial appraisal of what construct should be at the center of an intervention needs to be conducted before implementing any intervention in educational practice. In some cases, the total effect of an intervention that is known to target a different construct but that has an indirect effect on self-concept may also be the most attractive choice. For instance, attributional retraining (see Försterling, 1985) seems to be an effective intervention that also has a positive impact on academic self-concept (also see Hattie, 1992). Similarly, some recent examples of powerful interventions are described in Chap. 10 of this volume.

8.8 Conclusion

Academic self-concept is an important outcome of schooling and a powerful predictor of students' academic behavior, educational choices, and academic achievement. Given its theoretical and practical importance, it's not surprising that it has become one of the most prominent constructs in educational psychology. There are a number of different sources that impact the development of students' academic self-concept, including social, dimensional, and temporal comparisons, and the study of the complex pattern of influences on self-concept remains a very active research field. There is some (but not complete) consensus in the literature that schools should foster relatively high levels of self-concept. However, given the seemingly universal nature of social and dimensional comparisons and empirical data showing comparably minor differences across classes, schools, and school types, the improvement of academic self-concept does seem to be a rather challenging task. For this reason—and despite the powerful effects of academic self-concept on a number of important academic outcomes—self-concept may not always be the most meaningful target of educational interventions. In addition, more research is needed that examines intervention effects of expectancy-value intervention studies on self-concept.

Acknowledgment Some parts of this chapter were informed by the authors' chapter on self-concept in a German language textbook (Möller & Trautwein, 2015).

References

Albert, S. (1977). Temporal comparison theory. *Psychological Review, 84*, 485–503.

Archambault, I., Eccles, J. S., & Vida, M. N. (2010). Ability self-concepts and subjective value in literacy: Joint trajectories from grade 1 through 2. *Journal of Educations Psychology, 102*(4), 804–816.

Asendorpf, J. B., & van Aken, M. A. G. (2003). Personality-relationship transaction in adolescence: Core versus surface personality characteristics. *Journal of Personality, 71*, 629–662.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Baumeister, R. F., Campbell, J. C., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest, 4*, 1–44.

Bong, M. (2001). Between- and within-domain relations of academic motivation among middle and high school students: Self-efficacy, task value, and achievement goals. *Journal of Educational Psychology, 93*, 23–34.

Brunner, M., Keller, U., Dierendonck, C., Reichert, M., Ugen, S., Fischbach, A., et al. (2010). The structure of academic self-concepts revisited: The nested Marsh/Shavelson model. *Journal of Educational Psychology, 102*, 964–981.

Byrne, B. M. (1996). *Measuring self-concept across the life span: Issues and instrumentation*. Washington, DC: APA Books.

Calsyn, R. J., & Kenny, D. A. (1977). Self-concept of ability and perceived evaluation of others: Cause or effect of academic achievement? *Journal of Educational Psychology, 69*, 136–145.

Chiu, M.-S. (2012). The internal/external frame of reference model, big-fish-little-pond effect, and combined model for mathematics and science. *Journal of Educational Psychology, 104*, 87–107.

Chmielewski, K., Dumont, H., & Trautwein, U. (2013). Tracking effects depend on tracking type: An international comparison of academic self-concept. *American Educational Research Journal, 50*, 925–957.

Cialdini, R. B., & Richardson, K. D. (1980). Two indirect tactics of image management: Basking and blasting. *Journal of Personality and Social Psychology, 39*, 406–415.

Dai, D. Y., & Rinn, A. N. (2008). The big-fish-little-pond effect: What do we know and where do we go from here? *Educational Psychology Review, 20*, 283–317.

Deci, E. L., & Ryan, R. M. (2000). *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.

Denissen, J. J. A., Zarrett, N. R., & Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development, 78*, 430–447.

Diener, E., & Fujita, F. (1997). Social comparison and subjective well-being. In B. P. Buunk & F. X. Gibbons (Eds.), *Health, coping, and well-being: Perspectives from social comparison theory* (pp. 329–358). Mahwah, NJ: Erlbaum.

Dunlosky, J., & Rawson, K. (2012). Overconfidence produces underachievement: Inaccurate self-evaluations undermine students' learning and retention. *Learning and Instruction, 22*, 271–280.

Eccles, J. S. (1983). Expectancies, values, and academic choice: Origins and changes. In J. Spence (Ed.), *Achievement and achievement motivation* (pp. 87–134). San Francisco: W. H. Freeman.

Eccles, J.S., Wigfield, A., Harold, R.D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development, 64,* 830–847.

Eccles, J. S. (1994). Understanding women's educational and occupational choices. *Psychology of Women Quarterly, 11*, 135–172.

Eccles, J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., & Meece, J., et al. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives (pp. 75–146)*. San Francisco: W. H. Freeman.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*, 109–132.

European Commission. (2007). *Science education now: A renewed pedagogy for the future of Europe* (Community Research Expert Group, Vol. 22845). Luxembourg: EUR-OP.

Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction, 29*, 1–9.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117–140.

Försterling, F. (1985). Attributional retraining: A review. *Psychological Bulletin, 98*, 495–512.

Friedrich, A., Flunger, B., Nagengast, B., Jonkmann, K., Schmitz, B., & Trautwein, U. (2015). Pygmalion effects in the classroom: Teacher expectancy effects on students' math achievement. *Contemporary Educational Psychology, 41*, 1–12.

Frome, P. M., & Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of Personality and Social Psychology, 74*, 435–452.

Gogol, K., Brunner, M., Goetz, T., Martin, R., Ugen, S., Keller, U., et al. (2014). "My questionnaire is too long!" The assessments of motivational-affective constructs with three-item and single-item measures. *Contemporary Educational Psychology, 39*, 188–205.

Harter, S. (1998). The development of self-representations. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 553–617). New York: Wiley.

Hattie, J. (1992). *Self-concept*. Hillsdale, NJ: Lawrence Erlbaum.

Helmke, A. (1990). Mediating processes between children's self-concept of ability and mathematical achievement: A longitudinal study. In H. Mandl, E. de Corte, S. N. Bennett, & H. F. Friedrich (Eds.), *Learning and instruction. European research in an international context. Volume 2.2: Analysis of complex skills and complex knowledge domains* (pp. 537–549). Oxford: Pergamon Press.

Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology, 49*, 505–528.

Hulleman, C. S., Barron, K. E., Kosovich, J. J., & Lazowski, R. A. (2015). Student motivation: Current theories, constructs, and interventions within an expectancy-value framework. In A. A. Lipnevich, F. Preckel, & R. D. Roberts (Eds.), *Psychosocial skills and school systems in the 21st century*. Cham: Springer.

James, W. (1892/1999). The self. In R. F. Baumeister (Ed.), *The self in social psychology* (pp. 69–77). Philadelphia: Psychology Press. (Original: James, W. [1892/1948]. *Psychology*. Cleveland, OH: World Publishing).

Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review, 9*, 131–155.

Kim, Y.-H., Chiu, C.-Y., & Zou, Z. (2010). Know thyself: Misperceptions of actual performance undermine achievement motivation, future performance, and subjective well-being. *Journal of Personality and Social Psychology, 99*(3), 395–409.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121–1134.

Lüdtke, O., Köller, O., Marsh, H. W., & Trautwein, U. (2005). Teacher frame of reference and the big-fish-little-pond effect. *Contemporary Educational Psychology, 30*, 263–285.

Makel, M. C., Lee, S. Y., Olszewski-Kubilius, P., & Putallaz, M. (2012). Changing the pond, not the fish: Following high ability students across different educational environments. *Journal of Educational Psychology, 104*, 788–792.

Marsh, H.W., & Shavelson, R.J. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist, 20*, 107–125.

Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal, 23*, 129–149.

Marsh, H. W. (1987). The big fish little pond effect on academic self-concept. *Journal of Educational Psychology, 79*, 280–295.

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology, 81*, 417–430.

Marsh, H. W. (1990). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology, 82*, 623–636.

Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology, 80*, 366–380.

Marsh, H. W., Chessor, D., Craven, R., & Roche, L. (1995). The effects of gifted and talented programs on academic self-concept: The Big Fish strikes again. *American Educational Research Journal, 32*, 285–319.

Marsh, H. W., & Craven, R. (1997). Academic self-concept: Beyond the dustbowl. In G. D. Phye (Ed.), *Handbook of classroom assessment* (pp. 137–198). San Diego, CA: Academic.

Marsh, H. W., & Craven, R. (2002). The pivotal role of frames of reference in academic self-concept formation: The big-fish-little-pond effect. In F. Pajares & T. Urdan (Eds.), *Adolescence and education* (Vol. 2, pp. 83–123). Greenwich, CT: Information Age.

Marsh, H. W., Craven, R., & Debus, R. (1998). Structure, stability, and development of young children's self-concepts: A multicohort-multioccasion study. *Child Development, 69*, 1030–1053.

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science, 1*, 133–162.

Marsh, H. W., & Hattie, J. (1996). Theoretical perspectives on the structure of self-concept. In B. A. Bracken (Ed.), *Handbook of self-concept* (pp. 38–90). New York: Wiley.

Marsh, H. W., & Hau, K.-T. (2003). Big-Fish-Little-Pond effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist, 58*, 364–376.

Marsh, H. W., Kong, C.-K., & Hau, K.-T. (2000). Longitudinal multilevel models of the big fish little pond effect on academic self-concept: Counterbalancing contrast and reflected-glory effects in Hong Kong schools. *Journal of Personality and Social Psychology, 78*, 337–349.

Marsh, H. W., Kong, C.-K., & Hau, K.-T. (2001). Extension of the internal/external frame of reference model of self-concept formation: Importance of native and nonnative languages for Chinese students. *Journal of Educational Psychology, 93*, 543–553.

Marsh, H. W., Kuyper, H., Seaton, M., Parker, P. D., Morin, A. J. S., Möller, J., et al. (2014). Dimensional comparison theory: An extension of the internal/external frame of reference effect on academic self-concept formation. *Contemporary Educational Psychology, 39*, 326–341.

Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Abduljabbar, A.S., Abdelfattah, F., et al. (2015). Dimensional comparison theory: Paradoxical relations between self-beliefs and achievements in multiple domains. *Learning and Instruction, 35*, 16–32.

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology, 81*, 59–77.

Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J., & Köller, O. (2007). Big Fish Little Pond Effect: Persistent negative effects of selective high schools on self-concept after graduation. *American Educational Research Journal, 44*, 631–669.

Marsh, H. W., Trautwein, U., Lüdtke, O., & Köller, O. (2008). Social comparison and big-fish-little-pond effects on self-concept and efficacy perceptions: Role of generalized and specific others. *Journal of Educational Psychology, 100*, 510–524.

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*, 397–416.

Marsh, H. W., & Yeung, A. S. (1997). Coursework selection: Relations to academic self-concept and achievement. *American Educational Research Journal, 34*, 691–720.

Marsh, H. W., & Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: Gender differences in the development of math and English constructs. *American Educational Research Journal, 35*, 705–738.

Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology, 82*, 60–70.

Möller, J., Helm, F., Müller-Kalthoff, H., Nagy, N., & Marsh, H. W. (2015). Dimensional comparisons: Theoretical assumptions and empirical results. In J. D. Wright (Ed.), *International encyclopedia of the social and behavioral sciences* (2nd ed., pp. 430–436). Oxford: Elsevier.

Möller, J., & Husemann, N. (2006). Internal comparisons in everyday life. *Journal of Educational Psychology, 98*, 342–353.
Möller, J., & Köller, O. (2001). Dimensional comparisons: An experimental approach to the internal/external frame of reference model. *Journal of Educational Psychology, 93*, 826–835.
Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review, 120*, 544–560.
Möller, J., & Pohlmann, B. (2010). Achievement differences and self-concept differences: Stronger associations for above or below average students? *British Journal of Educational Psychology, 80*, 435–450.
Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research, 79*, 1129–1167.
Möller, J., Retelsdorf, J., Köller, O., & Marsh, H. W. (2011). The reciprocal I/E model: An integration of models of relations between academic achievement and self-concept. *American Educational Research Journal, 48*, 1315–1346.
Möller, J., & Savyon, K. (2003). Not very smart thus moral: Dimensional comparisons between academic self-concept and honesty. *Social Psychology of Education, 6*, 95–106.
Möller, J., Zimmermann, F., & Köller, O. (2014). The reciprocal internal/external frame of reference model using grades and test scores. *British Journal of Educational Psychology, 84*, 591–611.
Nagengast, B., Marsh, H.W., Scalas, L.F., Xu, M., Hau, K.-T., & Trautwein, U. (2011). Who took the "X" out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science, 22*, 1058–1066.
Nagy, G., Trautwein, U., Baumert, J., Köller, O., & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation, 12*, 323–345.
Niepel, C., Brunner, M., & Preckel, F. (2014). The longitudinal interplay of students' academic self-concepts and achievements within and across domains: Replicating and extending the reciprocal internal/external frame of reference model. *Journal of Educational Psychology, 106*, 1170–1191.
OECD (Ed.). (2001). Knowledge and skills for life: First results from PISA 2000. Paris: OECD.
O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. (2006). Do self-concept interventions make a difference? A synergetic blend of construct validation and meta-analysis. *Educational Psychologist, 41*, 181–206.
Oakes, J. (1985). *Keeping tack: How schools structure inequality*. New Haven, CT: Yale University Press.
Pianta, R. C., & Hamre, B. K. (2009). Classroom processes and positive youth development: Conceptualizing, measuring, and improving the capacity of interactions between teachers and students. *New Directions for Youth Development, 121*, 33–46.
Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology, 95*, 667–686.
Pinxten, M., Wouters, S., Preckel, F., Niepel, C., De Fraine, B., & Verschueren, K. (2015). The formation of academic self-concept in elementary education: A unifying model for external and internal comparisons. *Contemporary Educational Psychology, 41*, 124–132.
Preckel, F., & Brüll, M. (2010). The benefit of being a big fish in a big pond: Contrast and assimilation effects on academic self-concept. *Learning and Individual Differences, 20*, 522–531. doi:10.1016/j.lindif.2009.12.007.
Retelsdorf, J., Köller, O., & Möller, J. (2014). Reading achievement and reading self-concept: Testing the reciprocal effects model. *Learning and Instruction, 29*, 21–30.
Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom*. New York: Holt, Rinehart & Winston.
Ross, M., & Wilson, A. E. (2003). Autobiographical memory and conceptions of self: Getting better all the time. *Current Directions in Psychological Science, 12*, 66–69.

Ryan, R. M., & Weinstein, N. (2009). Undermining quality teaching and learning: A self-determination theory perspective on high-stakes testing. *Theory and Research in Education, 7*, 224–233.

Schwarzer, R., Lange, B., & Jerusalem, M. (1982). Selbstkonzeptentwicklung nach einem Bezugsgruppenwechsel. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 14*, 125–140.

Seaton, M., Marsh, H. W., Duman, F., Huguet, P., Monteil., J. M., Régner, I., et al. (2008). In search oft he big fish: Investigating the coexistence of the big-fish-little-pond effect with the positive effects of upward comparisons. *British Journal of Social Psychology, 47*, 73–103.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Validation of construct interpretations. *Review of Educational Research, 46*, 407–441.

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology, 42*, 70–83.

Suls, J., & Wheeler, L. (2000). *Handbook of social comparison: Theory and research*. New York: Kluwer Academic/Plenum Press.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181–227). New York: Academic.

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology, 92*, 144–151.

Tiedemann, J. (2002). Gender-related beliefs of teachers in elementary school mathematics. *Educational Studies in Mathematics, 41*, 191–207.

Trautwein, U., Köller, O., Lüdtke, O., & Baumert, J. (2005). Student tracking and the powerful effects of opt-in courses on self-concept: Reflected-glory effects do exist after all. In H. W. Marsh, R. G. Craven, & M. McInerney (Eds.), *New frontiers for self research* (International advances in self research, Vol. 2, pp. 307–327). Greenwich, CT: Information Age Press.

Trautwein, U., & Lüdtke, O. (2007). Students' self-reported effort and time on homework in six school subjects: Between-student differences and within-student variation. *Journal of Educational Psychology, 99*, 432–444.

Trautwein, U., Lüdtke, O., Kastens, C., & Köller, O. (2006). Effort on homework in grades 5 through 9: Development, motivational antecedents, and the association with effort on classwork. *Child Development, 77*, 1094–1111.

Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Self-esteem, academic self-concept, and achievement: How the learning environment moderates the dynamics of self-concept. *Journal of Personality and Social Psychology, 90*, 334–349.

Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth grade mathematics. *Journal of Educational Psychology, 98*, 788–806.

Trautwein, U., Lüdtke, O., Marsh, H. W., & Nagy, G. (2009). Within-school social comparison: How students' perceived standing of their class predicts academic self-concept. *Journal of Educational Psychology, 101*, 853–866.

Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Journal of Educational Psychology, 104*, 763–777.

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relations between self-beliefs and academic achievement: A systematic review. *Educational Psychologist, 39*, 111–133.

Wakslak, C. J., Nussbaum, S., Liberman, N., & Trope, Y. (2008). Representations of the self in the near and distant future. *Journal of Personality and Social Psychology, 95*, 757–773.

Watt, H. M. G., & Eccles, J. S. (Eds.). (2008). *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences*. Washington, DC: APA Books.

Wheeler, L., & Suls, J. (2005). Social comparison and self-evaluation of competence. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 566–578). New York: Guilford.

Wheeler, L., & Suls, J. (2007). Assimilation in social comparison: Can we agree on what it is? *Revue Internationale de Psychologie Sociale, 20*, 31–51.

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review, 12*, 265–310.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of motivation. *Contemporary Educational Psychology, 25*, 68–81.

Wigfield, A., & Eccles, J. S. (2002). Children's motivation during the middle school years. In J. Aronson (Ed.), *Improving academic achievement: Contributions of social psychology*. San Diego, CA: Academic.

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A., Freedman-Doan, K., et al. (1997). Changes in childrens' competence beliefs and subjective task values across the elementary school years: A three-year study. *Journal of Educational Psychology, 89*, 451–469.

Wigfield, A., Tonks, S., & Klauda, S. L. (2009). Expectancy – value theory. In K. R. Wentzel & A. Wigfield (Eds.), Handbook of motivation at school (pp. 55-76). New York: Taylor Francis.

Wilson, A. E., & Ross, M. (2000). The frequency of temporal-self and social comparisons in people's personal appraisals. *Journal of Personality and Social Psychology, 78*, 928–942.

Wilson, A. E., & Ross, M. (2001). From chump to champ: People's appraisals of their earlier and present selves. *Journal of Personality and Social Psychology, 80*, 572–584.

Zell, E., & Alicke, M. D. (2009). Self-evaluative effects of temporal and social comparison. *Journal of Experimental Social Psychology, 45*, 223–227.

Chapter 9
Applying Social Cognitive Theory in the Development of Self-Regulated Competencies Throughout K-12 Grades

Héfer Bembenutty, Marie C. White, and Maria K. DiBenedetto

9.1 Applying Social Cognitive Theory in the Development of Self-Regulated Competencies Throughout K-12 Grades

Self-regulated learning refers to the processes used by learners to systematically and actively attain their personal goals. As such, it involves an individual's use of cognition, behavior, and affect to sustain this pursuit (Zimmerman & Schunk, 2011). Bandura's (2012) social cognitive theory explains individual learning, development, acquisition of knowledge, and self-regulated competency within a social context, in which parents, peers, and teachers play a significant role as social models. Social cognitive theory research has been of interest to educators, parents, and policymakers because of its explanatory power in understanding human development, its practicality, and its applicability to learning. A unique contribution of social cognitive theory is its emphasis on the triadic interaction between the person, behavior, and the environment. This triadic model accounts for how individuals are not dependent exclusively on the environment but are able to manipulate, react, and influence their environment. Personal factors such as cognition and affect also come into play.

H. Bembenutty (✉)
Department of Secondary Education and Youth Services, Queens College of the City University of New York, 65-30 Kissena Blvd, Queens, NY 11367-1597, USA
e-mail: hefer.bembenutty@qc.cuny.edu

M.C. White
School of Education, Nyack College NYC Campus,
2 Washington Street, New York, NY, USA
e-mail: Marie.White@nyack.edu

M.K. DiBenedetto
Bishop McGuinness Catholic High School, 1725 HW 66 S, Kernersville, NC, USA
e-mail: mkdibenedettophd@aol.com

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_9

One of the main assumptions underlying social cognitive theory is that individuals are competent and active agents whose actions can influence their development, learning, and behavior (Bandura, 1997). Bandura (1997) suggests that individuals are able to learn by observing others in social contexts. He posits that individuals respond to their environment based on their beliefs, values, prior experiences, sense of efficacy, and expectancies. Social cognitive theory also conceptualizes individuals as able to engage in the regulation of their own thoughts, beliefs, and actions.

Under the umbrella of social cognitive theory, research on self-regulated learning has attracted a significant amount of attention during the last four decades. Self-regulated learning has practical applications to the classroom context and instruction that are far reaching. Several studies place self-regulation as the cornerstone of academic endeavors and consider it a determinant of learning and development (Bembenutty, Cleary, & Kitsantas, 2013; Zimmerman & Schunk, 2011). This line of research emphasizes the role of teachers as important social models and the role of learners as proactive and self-directed seekers in all learning endeavors. In this vein, the role of self-efficacy is conceptualized as a determinant of actions and an agent of learning, and self-regulation is construed as a process of growth and development (see e.g., Bembenutty et al., 2013; Boekaerts, Pintrich, & Zeidner, 2000; Zimmerman & Schunk, 2011).

In the early years of the development of social cognitive theory, Bandura conducted seminal work on self-regulation and self-efficacy (Bandura, 1986). Whereas *self-regulated learning* refers to individuals' control of their thoughts, actions, feelings, and behavior in order to pursue designated academic tasks (Zimmerman, 2013), *self-efficacy* is defined as individuals' beliefs about their capability to perform designated tasks. Consistent with Bandura's theory, Zimmerman's research on self-regulation reveals that learners are able to identify learning targets toward which they set goals, select strategies for learning and monitoring, and are able to engage in self-reflection of outcomes.

Zimmerman (e.g., 2000; Zimmerman & Kitsantas, 1997; Schunk & Zimmerman, 1997) suggests that learners develop self-regulation as they move through four levels. The first level begins by observing an effective and competent model followed by the second level, which involves emulating the patterns observed under the direct supervision of the model. Once learners have the observed patterns of behavior and are able to replicate these patterns, they engage in the process of self-control, in which they perform tasks somewhat independent of the model while keeping the initial represented pattern in mind. When learners are able to do the task independently of the model, they are able to adapt and modify the initial techniques as needed by the context. This final level is called self-regulation. However, there is limited understanding of how self-regulated learning can be applied to different developmental stages of children in the classroom. For example, it appears self-evident that kindergarteners will process information differently than high school seniors. Teachers and peers also play significant, but different, roles at various maturational levels. How these different factors play out in the classroom appears not so well understood.

To address this gap, the first section of this chapter reviews the research on social cognitive theory with particular emphasis on the role of modeling, self-efficacy, and self-regulated learning. The second section presents four models of self-regulated

learning and considers the role of self-regulated learning in various academic contexts with an emphasis on Zimmerman's developmental levels of acquiring self-regulated competency. Finally, the third section discusses developmental trends and applications of self-regulation in K-12 learning settings. We conclude by providing educational implications that will help to explain how learning happens in classroom settings, suggest classroom interventions for promoting self-regulated learning, and offer avenues for future research.

9.2 Social Cognitive Theory

Social cognitive theory provides a framework for learning that takes into account the social environment, the personal factors such as affect and cognition of the learner, and the behavior (Bandura, 1986, 2012). Consider Steven, a third grader with no prior history of school-related problems, who has become easily distracted, often forgets to do his homework, and exhibits outbursts during instruction time. As a result, his classmates often exclude him from games during recess and he does not seem to have any friends. Since Steven's performance in the second and first grades was without similar types of behavioral problems, his teacher has become increasingly concerned. Due to Steven's failure to complete his third-grade work, the teacher was worried about the upcoming state assessments. During the parent–teacher conferences, Steven's teacher discussed her concerns with his mom and learned that his parents had separated just before the start of the school year. Steven's mother stated that the boy was very upset about his father moving out and suspected that the drop in Steven's grades and increased behavior problems were related to the events taking place in the home. Steven's parents' separation and his father's leaving (home environment) appeared to be affecting his concentration (personal) and his performance (behavior) in school.

Bandura's (1986, 1997) theory suggests that there is reciprocity among the environmental, personal, and behavioral factors in this triadic model meaning that as they interact, they also determine or cause the other (see Fig. 9.1). Bandura's social cognitive theory is based on the assumptions that these three factors are influenced by enactive and observational learning from one's environment, personal motivation or self-efficacy, and the ability to self-regulate. Steven's behaviors are evidence of how a disruption in one area can impact performance in another.

9.2.1 Triadic Model

Humans have intrapersonal or innate characteristics such as biological strengths and weaknesses, values, and affective states that influence behavior (Bandura, 2007). Internal and external feedback from the environment, in turn, influences one's thoughts and affect. Reciprocity occurs when someone takes action and it has an impact on the environment. For example, when a student who is attempting to

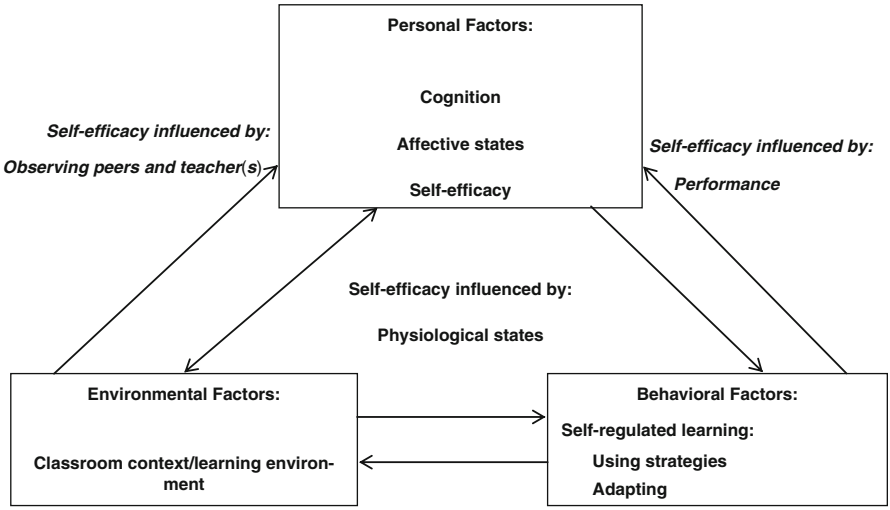

Fig. 9.1 Triadic model in a classroom context

understand statements made by the teacher during a lesson asks a question, the asking of the question can direct the teacher to slow down the pace to provide more details about the topic. Hence, in asking the question, the student's actions have altered the learning environment. The alteration in the teacher's behavior will aid in the student's understanding of the lesson. This exemplifies the triadic mode where behavior, personal, and environmental (or social) factors influence each other.

Bandura (2007) emphasizes the social aspect of the environment such as the classroom environment, teacher, and peers. Research studies have revealed the effects of one's social environment on achievement and competency beliefs (Bembenutty, 2008, 2011; Bembenutty et al., 2013; DiBenedetto & Bembenutty, 2011; Zimmerman & Schunk, 2011). In school settings, students are influenced by their peers and their teachers through observational learning as well as social persuasions (see Fig. 9.1; Bandura, 2007). In the above example, students who observe classmates asking questions and the teacher's positive manner of responding are likely to ask questions when they are confused. Social persuasions occur when teachers or peers encourage students to perform an activity. A student who is playing an instrument may be encouraged by his teacher to continue, even after playing the incorrect keys. This interpersonal exchange is part of the learner's environment and continually influences his or her thoughts, feelings, and behavior.

9.2.2 Observational Learning

In social cognitive theory, individuals can learn from making observations of others. Through these observations they learn many things, including how to perform a specific task. Through observational learning or modeling, a learner acquires new

patterns of behavior that may have had a zero probability of occurring prior to observing the model. Bandura (1986) emphasized four characteristics of models that make learners pay attention: prestige, competence, perceived similarity, and vicarious consequences for their actions. In a classroom setting, each of these characteristics may be present in the teacher(s) and in one's peers.

Teachers may enhance student's self-regulated learning by demonstrating problem-solving techniques and thinking aloud (i.e., verbalization of thoughts during problem-solving; Horner & Shwery, 2002; Zimmerman, 2013). Showing students how to persist in the face of failure by changing strategies to find the one that works is a significant feature of modeling. Research has demonstrated that observational learning, particularly in the form of cognitive modeling, is a powerful tool for teachers to use in the classroom (Zimmerman, 2013). When teachers use cognitive modeling, they not only demonstrate *what to do* but *how to think about* the actions that need to be taken to complete the task at hand successfully (Dorn & Sofos, 2001). This verbalization provides learners with the opportunity to observe the thinking patterns that lead to success (Schunk & Usher, 2012).

9.2.3 Self-Efficacy

Self-efficacy has been described as a belief one has about his or her capability to accomplish one's goals (such as achieving a high score on a math test, playing a concerto perfectly, or shooting a basketball into a hoop). Self-efficacy has been demonstrated to relate to one's efforts, persistence, and choice of activity (e.g., Bandura, 1997; Pajares, 1996; see also Chaps. 8 and 10). For example, young gymnasts who are self-efficacious about their ability to flawlessly walk across an elevated beam are likely to consistently practice and not give up, even after falling, until they are successful. Self-efficacy develops as a result of four factors: social persuasion, physiological responses, observations of others, and mastery experiences (Bandura, 1997).

Social persuasion highlights the important role of how others can help shape a learner. When a learner receives praise or positive feedback, the source must be credible in order to raise self-efficacy (Bandura, 1986, 1997; Lipnevich & Smith, 2009; Schunk & DiBenedetto, 2015). In longitudinal research on music students, for example, young musicians' motivation to practice has been found to be positively linked to the frequency of mothers' practice reminders (McPherson & Davidson, 2002). Physiological and emotional states such as fear and anxiety also impact self-efficacy (Bandura, 1997). Students who worry about performing well on a test may experience agitation, nervousness, and stress that may interfere with their test performance. In other words, the anxiety and worry may impact learners' self-efficacy (Bandura, 1986, 1997; Schunk & DiBenedetto, 2014). Further, as indicated earlier, individuals learn vicariously. Witnessing the success of other learners who share similar characteristics may strengthen the self-efficacy of the viewers and motivate them to perform similar tasks (Bandura, 1997). Mastery experiences are critical for

building self-efficacy and research supports that mastery experiences result in higher levels of achievement (Zimmerman & DiBenedetto, 2008). Students who perform a task well are likely to be motivated to continue learning and their self-efficacy will be enhanced (Schunk & Mullen, 2013). Self-efficacy is not only critical for its personal motivational affects, but it is hypothesized to affect the behavior and environmental contexts within the triadic model (Schunk & Usher, 2012).

Students who feel self-efficacious are likely to set academic goals for themselves and use effective strategies to achieve these goals. They adapt and regulate their behavior as needed and accept responsibility for their own learning. Self-efficacy is considered a motivational process in that it is said to influence behavior (Schunk & DiBenedetto, 2014), as well as how much energy students will expend to complete various academic tasks (Schunk & Pajares, 2009). Self-efficacy is a pivotal process in self-regulated learning because students who are self-efficacious are more likely to engage in self-regulatory behaviors, whereas those who are not are less likely to feel motivated.

9.3 Models of Self-Regulated Learning

Important models of self-regulation have been developed during the last several years (Bembenutty et al., 2013; Boekaerts et al., 2000; Vohs & Baumeister, 2013; Zimmerman & Schunk, 2011). In this section we review four models of self-regulated learning that have been most commonly employed in educational research and instruction, those of (1) Winne, (2) Boekaerts, (3) Pintrich, and (4) Zimmerman. We begin by providing a brief overview of the first three models. Then, we focus extensively on Zimmerman's model because it has been applied widely in developmental research, is most applicable to the theme of this chapter, and because the model emphasizes the importance of social relationships in learning.

9.3.1 Winne's Model

Winne presents an information processing model of self-regulated learning (Winne & Perry, 2000). This model focuses on the learner's metacognition, defined as awareness of one's own strength and abilities in relation to the task demands, and specific ways in which learners cognitively adapt to the task demands to regulate their learning strategies. The model involves four distinct stages that are recursive in nature (Puustinen & Pulkkinen, 2001). During the first stage, learners process information that generates an understanding of what is needed for successful completion of the task (Winne, 2011). In this stage, learners examine the task in terms of external or internal constraints and engage in metacognition as they check their understanding and redefine the task as needed (task definition). In the second stage, learners are believed to set goals and plans for how to complete the task and use

metacognition to determine how to work on the task (goal setting and planning). In the third stage, learners work on the task at hand (Winne, 2011). Here, metacognition plays the role of monitoring the learning process by signaling to learners when they need to go back and review the material further or apply more effort (studying tactics) (Winne, 2001). In Winne's fourth and final stage, the learner reflects on the entire learning experience to determine what worked well and what needs to be changed (adaptations to metacognition). The main objective at this stage is to determine what changes need to be made in the future for similar tasks from feedback. The model is recursive in that output products that were created by information processing at any stage can jump back and forth or they can serve as future inputs when similar learning events arise.

9.3.2 Boekaerts' Model

Boekaerts' model of self-regulated learning looks at learning episodes, in which learners are asked to demonstrate context-specific, goal-directed learning behavior (Boekaerts & Niemivirta, 2000). An important point of this model is that learners should be encouraged to cocreate learning episodes that will help them to reach their own goals – as opposed to participating in episodes created exclusively by the teacher. Self-regulation involves being able to adapt to various learning episodes, including those set by the teacher. Boekaerts describes identification, interpretation, and appraisal processes as essential to a model of adaptable learning. Identification refers to the learners being able to recognize a learning event and all of the features associated with that event. For example, Boekaerts and Niemivirta (2000) suggest that if the teacher were to announce an upcoming quiz, the students would recognize the pattern of events of a test situation because they have encountered it before. Interpretation processes involve both being task focused and being self-focused. Being task focused occurs when learners draw inferences about whether an event will have desirable or undesirable consequences, whereas being self-focused takes into account the learners' motivational factors. Appraisals are considered to be central factors in this theory of self-regulated learning, and they are influenced by the identification and interpretation of the learning situation. Appraisals may be positively charged, leading to increased subject knowledge and skill or negatively charged, leading to ego protection and the loss of one's well-being (Puustinen & Pulkkinen, 2001). When there is a balance between these two types of appraisals, the learner is said to be self-regulated.

9.3.3 Pintrich's Model

Pintrich's model suggests that self-regulated learning occurs in four phases with processes falling into cognitive, motivational/affective, behavioral, and contextual areas (Pintrich, 2000). The first phase is forethought, planning, and activation.

During this phase learners set their goals, adopt a goal orientation, and determine the effort and time needed to complete the task. This phase also involves efficacy judgments and perceptions of task difficulty. The second phase is the monitoring phase. Here learners are engaging in metacognitive awareness and monitoring of various processes (e.g., cognition, motivation, and affect), of the effort and time needed to complete the task, and of the changing task and contextual conditions. In the third phase of control, learners select and adapt cognitive strategies for learning and thinking. They are managing motivation and affect, will increase or decrease effort, seek help, change or negotiate the task, and change or leave the learning context. In the final phase, learners are believed to react and reflect by forming cognitive judgments and attributions, choosing behavior, and evaluating the task and learning context.

Pintrich examined the role of motivation in students' mastery and performance orientations by examining students' approach versus avoidance viewpoint (Pintrich, 2000). Students who had a mastery-approach orientation were hypothesized to be more focused on learning as compared to students with a mastery-avoidance orientation where students were deemed to be more eager to escape imperfection. Students who took a performance-approach orientation were learners who were positively motivated to outperform others and to demonstrate their competence, whereas learners who took a performance avoidance orientation were motivated to avoid failure and looking incompetent to their peers. From the self-regulated learning perspective, students who engage in a mastery approach orientation are more likely to show positive academic results (Puustinen & Pulkkinen, 2001).

9.3.4 Zimmerman's Model

Zimmerman (2000) proposed a three-phase model of self-regulation, in which learners engage in a cyclical self-regulated learning process. Self-regulation of learning is cyclically initiated when learners set valuable academic goals, select learning strategies, and assess their feelings and motivational beliefs necessary to attain the goals. In Zimmerman's model, students establish standards, monitor their academic progress, and evaluate their academic progress. During the *forethought* phase, learners generate goals, identify learning strategies, assess their intrinsic and extrinsic motivation, and assess their degree of self-efficacy beliefs. During the *performance phase*, learners monitor their goals, beliefs, and use of strategies by comparing their performance with appropriate standards. Learners also seek help from appropriate knowledgeable sources, such as parents, teachers, peers, and technological resources while engaging in social and environmental control. During the *self-reflective* phase, learners evaluate their academic progress and examine their level of satisfaction with the completed task. If the outcome of the self-reflection is positive, learners move on to the next task; otherwise, they go to the previous phases to try to correct errors. For instance, learners who attribute unsatisfied outcomes to

use of incorrect learning strategy could go back and start again by using a different strategy. Zimmerman advanced another self-regulated model with an emphasis on the internal processes of the development of self-regulatory competence. The next section highlights the patterns of the developmental model of self-regulation.

9.4 The Development of Self-Regulatory Competence

Zimmerman's (2000) theory stems from the social cognitive perspective that views learning as initiating from outside or from external sources and shifting to internal sources (Schunk & Zimmerman, 1997). Zimmerman (2000) defined self-regulated learning as one's self-generated feelings, thoughts, and behaviors geared toward attaining one's goals. Social cognitive theory suggests that modeling, which occurs outside of the observer in the environment, is a precursor to self-regulation (Schunk & Zimmerman, 1997; Zimmerman, 2000; Zimmerman & Kitsantas, 1997). Self-efficacy, which begins externally by means of modeling, ultimately becomes internalized when self-regulation is reached (Schunk & Zimmerman, 1997; Zimmerman, 2000). Initially, this internalization appears socially, for instance, when students observe teachers and peers. Through practice and mastery experiences, students are more self-efficacious and competent, and they depend less on others' as they internalize feelings of self-efficacy and become self-directed learners (DiBenedetto & White, 2013; Zimmerman & DiBenedetto, 2008). Research suggests that several of the processes in Zimmerman's model of self-regulated learning have been linked to academic achievement such as strategic planning, metacognitive monitoring, and self-reflection (DiBenedetto & Zimmerman, 2010). Self-regulation is dynamic because learners continually evaluate and adapt their performance against a standard or a model (DiBenedetto & White, 2013). As seen in Fig. 9.2, self-regulation is developed through four sequential levels: observation, emulation, self-control, and self-regulation (Schunk & Zimmerman, 1997; Zimmerman, 2000; Zimmerman & Kitsantas, 1997).

Level 1: Observation Through observations of a more competent model, students learn about many self-regulatory processes such as how to plan for learning strategies, how to self-monitor performance, or how to set goals (Bandura, 1986; Schunk & DiBenedetto, 2015; Zimmerman, 2000). In addition, students' observation of models can include receiving rewards such as successful completion of a task or praise from a teacher. These observations enhance students' beliefs in their own capabilities to perform the observed behavior, thus initiating self-efficacy. At this level, students are observing the teacher who conducts cognitive modeling as she explains, demonstrates, and verbalizes reasons for her actions (Pape, Bell, & Yetkin-Ozdemir, 2013). The teacher controls the learning environment as she paces her instruction according to the students' readiness to learn the course content (DiBenedetto & White, 2013), although often knowledgeable peers could also serve as models and may impart information.

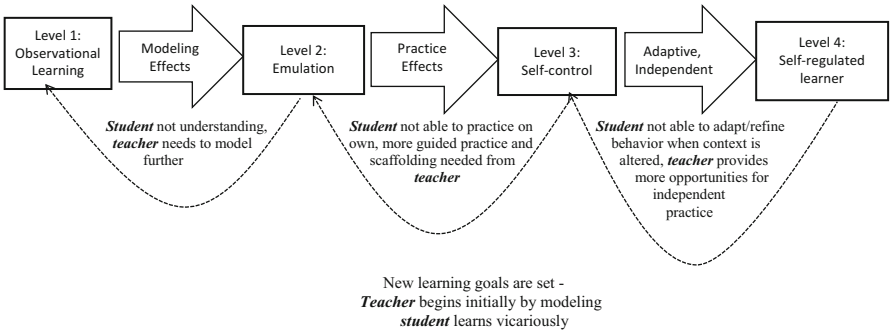

Fig. 9.2 Four levels of development of self-regulatory competency within a social learning context

Level 2: Emulation At this level students practice the behavior previously observed by the teacher and reproduce the general pattern or style with assistance from the model (Zimmerman, 2000). The teacher reinforces the students, and they begin to experience a sense of mastery as they closely replicate their observations of the teacher's cognitive modeling and observed behavior (DiBenedetto & White, 2013). Teachers provide very specific learning tasks that students are able to emulate in addition to providing them with frequent feedback (Pape et al., 2013). The teacher paces instruction to accommodate the students, but at this level students are beginning to internalize the self-regulatory processes and gain self-efficacy (DiBenedetto & White, 2013). As seen in Fig. 9.2, if students are not able to understand or replicate the patterns of modeled behavior, the teacher returns to level 1 and conducts further demonstrations to meet the needs of the students (Schunk & Zimmerman, 1997).

Research supports the link between achievement and the sequential development of these two levels. For example, in a study conducted on fifth graders learning math, students who observed models using problem strategies performed better than those who did not during the emulation level (Ramdass, 2011). In addition, similar results were found on an athletic task among high school girls and a writing task among college students (Kitsantas, Zimmerman, & Cleary, 2000; Zimmerman & Kitsantas, 2002) hence providing further evidence for the development of self-regulatory competency.

Level 3: Self-Control The third level of development of self-regulatory competency is that of self-control. Self-control initiates when students begin to experience self-efficacy from within. Students at the self-control level are able to use strategies independently of the presence of a model. Students have internalized what they have previously observed and emulated but are still using representational patterns of their model in their performance behaviors. At this level, students self-reinforce and reflect upon their capability of matching their work against the standards of the model. The source of self-efficacy and self-regulated behavior no longer comes

from the teacher but is now within the learner (Schunk & Zimmerman, 1997; Zimmerman, 2000). It is here that the control of the pacing learning shifts from teacher to students (DiBenedetto & White, 2013).

At the self-control level, the student is analyzing tasks, setting goals, and working on tasks with little guidance from the teacher (Pape et al., 2013). The teacher's role shifts now to allowing students to do work independently, providing feedback, and adjusting the tasks when students experience difficulty (Pape et al., 2013). As seen in Fig. 9.2, when students are not able to exercise self-control in working independently, the teacher provides them with more opportunities for guidance and scaffolding.

Level 4: Self-Regulation Although the previous three levels involve the development of self-regulation, in actuality, the process of self-regulation is achieved at the fourth level. Students reach self-regulation when they are able to systematically adapt their performance to different conditions and situations (Zimmerman, 2000). At this level, students are motivated by their own self-efficacy beliefs. They are now capable of initiating the use of strategies, making adjustments to behavior based on the situation, and evaluating their performance with the understanding that they are capable of competently making changes as needed. Students at this level are able to regulate their learning without the assistance of a teacher. This level is marked by independence; the learner is motivated to achieve, will set his or her goals, will make adjustments depending on the situation, and is self-efficacious about his or her capability to perform (DiBenedetto & White, 2013; Schunk & Zimmerman, 1997).

In a recent study, high school juniors who engaged in self-regulatory processes performed better than those who did not when studying and taking a test on tornados (DiBenedetto & Zimmerman, 2010). Research on self-regulated learning and mathematics has been conducted in samples of younger students (Ramdass, 2011) through college students (Zimmerman & Schunk, 2011). This body of research provides evidence of the link between self-regulated learning and student academic attainment. While evidence exists on the relationship between self-regulated learning and achievement, how does one go about implementing self-regulated learning in authentic learning contexts across different grade levels? The following section will provide examples of how self-regulated learning may be used in the classroom to promote the development of self-regulated learning competency.

9.5 Developmental Trends, Self-Regulatory Processes, and Applications in K-12 Learning Settings

Classroom instruction must be developmentally appropriate, taking into account the physical, cognitive, social, and emotional characteristics and abilities that a particular age-group is likely to have (Ormrod, 2013), even with the diverse populations of the twenty-first century learning environments. Over time, students must increase

their awareness of the demands of academic tasks; practice consistent monitoring of their behavior, thoughts, and strategies; consider whether or not they are meeting the demands of diverse academic tasks and learning environments; and demonstrate growth in the ability to change aspects of a current level of performance when needed to accomplish a goal. Various positive outcomes for children and adolescents who are self-regulated include not only better academic performance but evidence of stronger problem-solving skills and higher levels of reading comprehension. The development of self-regulated learners across age-groups requires a learning environment where students focus on personal progress and view errors as learning opportunities (Perry & VandeKamp, 2000). Students are more likely to become self-regulated learners when their teachers serve as effective models that closely monitor how the social context influences learning and guide students toward becoming proactive learners (Lombaerts, Engels, & van Braak, 2008). Numerous intervention studies and meta-analyses show that intervention programs aimed at enhancing students' self-regulated learning can be designed and successfully implemented as part of classroom instruction (for reviews of interventions and their effectiveness see Dignath, Buettner, & Langfeldt, 2008; Paris & Paris, 2001; Schunk & Ertmer, 2000).

Teachers and students have distinct roles and specific responsibilities in fostering the development of self-regulation. In this section, we discuss the role of the teacher and the student at each developmental level of self-regulation by combining them into grade levels using developmental trends and age level characteristics, K-2, 3–5, 6–8, and 9–12. We acknowledge the diversity in development and understand that grade or age only serves as guidelines for identification for a specific population of learners. Keeping in mind that children do not always choose their environment, they actively and intentionally think about and act on their environments, and in doing so, they alter their environments in significant ways. Our discussion distinguishes the role of the teachers and the students in the process of developing goal setting, self-efficacy, help seeking, and self-evaluation as critical components of self-regulated learning at each developmental level in diverse learning contexts.

9.5.1 Developmental Trends: K-2 Grades

Children in the early grades (K-2) enter school having learned important life skills vicariously from the adults and peers in their social environment. They are beginning to distinguish between their own and others' perspectives, recognizing that their own thoughts and feelings might be different from others (Ormrod, 2013). Their development of cognitive self-regulation is highly influenced by an environment that provides them a support system where the behaviors of adults significantly influence their ability to carry out cognitive activities independently (Zimmerman & Schunk, 2011). Modeling of ordinary activities has already provided a foundation for observational learning. Self-regulation skills can take time to

develop in young learners, so it is important that teachers have developmentally appropriate expectations when presenting new information (Perry et al., 2002). As their working memory functions develop, children can successfully participate in activities that require following complex directions (Bronson, 2000).

Developmental characteristics of K-2 learners include children's ability to recognize simple thoughts in themselves and others and to internalize a model's standards for behavior at a basic level. At this level of development, children demonstrate an emerging ability to choose and set goals for learning and achievement. One important aspect of this developmental level is the ability to differentiate between appropriate and inappropriate behaviors and actions.

Consistent with Zimmerman's developmental model of self-regulation, during observation, the teacher keeps in mind the short attention span of the young observers and is careful when demonstrating how to set short, proximal goals for an upcoming task. In return, the learners are able to pay close attention to the model for short, incremental demonstrations of goal setting for a specific task, deciding on how many characteristics should be identified and retained. Similarly, during emulation, the teacher closely monitors how well the learners imitate the model, asking students to state the steps required to set short, proximal goals for this particular task. At the self-control level, learners begin to gain independence as they practice setting goals for a task. As they continue to emulate the model, the learners do not become easily discouraged because their set goals are within their ability and developmentally appropriate. During self-control, the teacher slowly withdraws her influence but continues to monitor individual progress maintaining oversight and calling attention to specific areas that need improvement. When students reach the self-regulation level, the teacher assigns work that will require the learners to utilize the newly acquired skills when no longer under the supervision of the teacher. Helping their students move beyond self-control to self-regulation is one practice of highly effective teachers.

9.5.2 Developmental Trends: Third to Fifth Grades

Third through fifth graders enter the classroom having been exposed to various types of models and vicarious learning experiences. Earlier impressions about performance, ability, and efficacy tend to influence motivation to participate in a learning environment more challenging than the first 3 years of formal schooling. Their ability to logically reason about abstract and hypothetical situations helps them to separate strategies for effective and ineffective ways of completing tasks in an academic setting.

Developmental trends at the third- to fifth-grade levels include children's abilities to select strategies, focus on tasks, identify goals, and master the use of some self-regulatory strategies. According to Ormrod (2011), children's at this grade level improve in their ability to assess their performance and persist on difficult tasks. One important aspect of this developmental level is the ability to initiate some

level of help seeking as a self-regulatory strategy to complete difficult tasks. Help seeking refers to strategic self-regulatory behaviors that learners use to obtain information from formal and informal sources to adapt and acquire knowledge and skills (Karabenick & Berger, 2013). At this developmental level, help seeking can be successfully integrated into teaching and learning practices particularly because children are intrinsically motivated to learn and are less competitive as compared to older ages (White & DiBenedetto, 2015). As children progress to middle school, they find it more difficult to seek assistance unless they have been provided with models of the benefits of help seeking in the earlier grades. As they get older and face challenging or difficult school work, even though they are aware that assistance is available, they will give up prematurely, sit passively, or persist unsuccessfully on their own without asking for help (Newman, 2000).

By modeling and then integrating training in help-seeking strategies into a challenging task, teachers can turn potential failures into learning opportunities (Graham & Harris, 1989; Troia & Graham, 2002). At the observation level, it is important for the student to observe the teacher demonstrating how and why asking for help with a difficult assignment leads to a successful completion of a task. This encourages the learner to pay attention and practice the same strategy. At the emulation level and under strict monitoring of the teacher, learners can practice help seeking and use positive and instructive feedback and at the same time complete a very challenging task. However, at the self-control level, the teacher is responsible to not only direct students' requests for assistance from competent models and peers but to also acquaint learners with a variety of resources that can be accessed somewhat independently during the self-control level. Finally, at the self-regulation level and once students have experienced the benefits of asking and receiving help, they reach the level of self-regulation, in which they apply the newly acquired strategy in other settings independent of the teacher.

9.5.3 Developmental Trends: Sixth to Eighth Grades

Students in grades six to eight are developmentally ready to master some self-regulatory learning strategies, especially those frequently observed. Many look for opportunities to learn independently and respond to explicit guidance on how to study effectively. However, changes in academic learning environments in middle school can lower students' self-efficacy and decrease their motivation when they transition from being in one classroom with the same peers and teacher for all subjects to several different classrooms with a mix of peer and teachers. Students build their efficacy beliefs through the vicarious experience of observing others; consequently, in many academic endeavors students gauge their capabilities in relation to the performance of others in the context of their learning environment (Schunk, 1987).

The developmental progression toward self-regulation might give the appearance of being fairly straightforward; however, educators have come to learn that each of

the levels of self-regulatory development is also dependent on contextual factors. Students construct knowledge through prior academic (and other) experiences. These self-beliefs mediate between the capacity to perform a skill and performance outcomes (Garcia & Pintrich, 1996).

Specific to self-regulation of learning, researchers have suggested that models play influential role during the shift from elementary to middle school, which in time students become more attuned to what others are doing (Eccles, Midgley, & Adler, 1984). Watching a classmate succeed at a challenging task, for instance, may build self-efficacy and convince fellow students that they too can conquer the challenge (Usher & Pajares, 2006). Students can be taught the benefits of using more sophisticated and strategic behaviors when given a task that requires them to explore texts with unfamiliar vocabulary to answer specific research questions. Schunk and Rice (1993) found that students receiving remedial reading services benefitted from procedures that required extensive cognitive activity and informed them about strategy usefulness.

At the observation level, it is critical for the model to think aloud and identify areas of confusion that might arise while working on the task and to provide a detailed description and some possible solutions. While observing, the student is encouraged to identify the specific patterns of behavior, learning the skills or strategies being modeled. At the emulation level, the teacher provides guidelines and positive feedback to help students recognize areas of weakness and strength that emerge when doing the task. In turn, the student accepts and internalizes the teacher's assessment, asking for clarification of specific areas of confusion when needed. At the self-control level and after several sessions of guided practice, the student takes the initiative to attempt the task independently, relying on recent successful use of the strategy under the supervision of the teacher. At the self-regulated level, the student is confident that he or she has the necessary skills to independently complete tasks that require comprehension of difficult texts to respond to research questions. As a result, as learners increase successful task completion, their self-efficacy is raised (Bandura, 1986; Usher & Pajares, 2008).

9.5.4 Developmental Trends: Ninth to Twelfth Grades

Increased academic and social pressures accompany the transition from middle school to high school requiring students to make choices regarding how they respond to environmental influences. Giving into distractions from completing homework assignments has been known to interfere with academic success (White & Bembenutty, 2013). Students' inability to delay gratification is inversely related to homework completion, but homework is correlated with academic achievement, self-regulation, and self-efficacy (Bembenutty & Karabenick, 2004). It is important to consider how the homework process or any task done independently can be evaluated and enhanced through the lens of self-regulation. Students engage in goal setting and planning enact study tactics and strategies, implement monitoring, and

evaluate strategies (Lipnevich, MacCann, Krumm, Burrus, & Roberts, 2011). In addition, how students make adjustments based on their self-evaluations is important information for both the student and the teacher.

One of the methods used to promote student's self-regulation involves encouraging students to evaluate their own work. Studies on the effect of self-evaluation during learning have strongly indicated that students who do so outperform students who are not encouraged to self-evaluate their progress while completing a specific task (Zimmerman & Schunk, 2011). Zimmerman, Bonner, and Kovach (1996) introduced the use of homework logs as an effective means for self-monitoring. This technique has been effectively used by elementary, middle school, high school, and college level students (Bembenutty, 2013; Bembenutty & White, 2013; Cooper, Horn, & Strahan, 2005). To illustrate, an intervention program designed by Cooper et al. (2005) used Zimmerman's model of self-regulatory training with students engaged in homework completion. Findings indicated that although at times the homework logs required a significant amount of attention, the logs helped students self-monitor, assess performance, and ultimately improve their grades. In order to use homework logs effectively, learners must be trained in self-recording behaviors and competences for a specific improved task.

It is more beneficial to learners if they can observe a teacher or peer demonstrate the steps of entering information into a homework log rather than just having the process explained (see Table 9.1). For instance, during observation, high school students' attention can be drawn to the model's verbalizations of how self-monitoring is beneficial in literacy learning tasks, citing past uses of the tool, and positive outcomes. At the emulation level, the teacher can closely monitor how well the students understand the value and usefulness of the homework log and explain how the process of self-evaluation can increase learning. While the teacher is modeling, the students begin to record the sequence of the steps and become more familiar with the benefits of self-evaluation. At the self-control level, the teacher prompts the students to use homework logs and shares with students how the logs can be used to ascertain the growth and development of specific skills. At the self-regulation level, the teacher challenges the students to use self-monitoring tools in other academic settings. Self-regulation is attained when the students independently begin to use homework logs and searches for other types of self-monitoring tools in areas that require close monitoring of academic growth and development.

9.6 Where Do We Go From Here?

In this section, we present educational implications and make suggestions for future directions of self-regulated learning research. The present discussion sought to expand one's understanding of educational practices, which could help learners acquire and maintain goal-directed learning even when faced with the temptation to do otherwise.

9.6.1 Educational Implications

We suggest that teacher preparation programs incorporate self-regulatory processes in their curriculum so that teacher candidates can use self-regulation in their own teaching as well as teach their students how to use these processes. It is in this training where teacher candidates will get their first glance of what their future careers will be. As Dembo (2001) observed, learning to teach is not enough; teacher candidates need to learn how to learn. Teacher preparation programs should educate future teachers to be social cognitive theorists who embrace observational learning, self-efficacy, and self-regulation.

Teachers and students need training on developing self-efficacy competencies. Students' inability to master academic tasks and teachers' low confidence in their capability to teach can be understood from the framework of self-efficacy. Teachers with high self-efficacy are those who believe that they can motivate even the most challenging student, can practice classroom management techniques, and can remain task focused. As Bandura (1997) has posited, schools are agencies for cultivating self-efficacy. Bandura (1997) stated, "during the crucial formative period of children's lives, the school functions as the primary setting for the cultivation and social validation of cognitive capabilities" (p. 174). He further argued that "a fundamental goal of education is to equip students with self-regulatory capabilities that enable them to educate themselves" (p. 174). Guided by the four developmental levels of self-regulation delineated by Zimmerman, teachers can find ways to build self-efficacy within themselves and among their students (Bembenutty, 2006; White, 2011; White & Bembenutty, 2013).

Educators and learners need to understand the importance of human agency in the triadic model. As Bandura (1997) suggested, the relationships between the person, the behavior, and the environment influence one another. Both teachers' and students' beliefs, affect, and cognition are influenced by their behavior and the environment in a reciprocal manner. In this way, in a classroom setting, both students and teachers are producers and products of this triadic reciprocality (Bandura, 1997).

In our discussion on self-regulation among students, particularly in grades 9–12, we emphasize that teachers should learn to appreciate and value the importance of delaying gratification of immediate rewards as a social and developmental process for the attainment of distant goals. Research supports the association between teachers who exhibit delay of gratification and positive classroom management, self-efficacy, intrinsic motivation, self-regulation, and teacher self-efficacy, all of which are considered effective components of teaching and instruction (Bembenutty, 2006; White & Bembenutty, 2013; see also Chap. 10). Similarly, teachers need to help their students to acquire the ability to delay gratification by modeling age-appropriate skills. From the perspective of social cognitive theory (Bandura & Mischel, 1965), we know that delay of gratification skills can be modeled and learned. A teacher could model to K-3 grade students on how to resist the urge to eat lunch until recess. Students in 4–5 grades can be introduced to the benefit of using weekly planners and homework logs. Students in grades 6–12 can be taught

Table 9.1 Developmental characteristics, grade levels, and applications of self-regulatory processes

Developmental characteristics and grade levels	Self-regulatory processes	Applications of the role of the teacher (T) and students (S)
K-2 Some ability to: Recognize simple ideas in oneself and others Internalize adult standards for behavior at basic level Self-evaluate appropriateness of actions Control impulses to sustain attention	Engage in goal setting to establish a standard or objective in order to pursue important tasks	Observation: T: Considers students' short attention span S: Focus attention on the teacher's modeling features Emulation: T: Closely monitors learners' imitation of model S: Attempt to imitate teacher's modeling behaviors Self-control: T: Remains available while closely monitoring students independent practice S: Set attainable goals using teacher feedback and guidance Self-regulation: T: Assigns work that requires independent practice S: Work independently
Grades 3–5 Improving ability to: Select strategies Focus on task Identify goals Master strategic use	Engage in help seeking to obtain information from appropriate sources in order to complete tasks	Observation: T: Demonstrates how to seek help when a task becomes too difficult S: Pay attention to the teacher's modeling features Emulation T: Closely monitors help-seeking behaviors of students to posed questions S: Imitate teacher's help-seeking behavior choosing appropriate resources Self-control: T: Slowly withdraws scaffolding while students begin to seek help S: Begin to formulate a plan to seek help using teacher feedback and guidance Self-regulation: T: Provides tasks that require students to seek help independently (homework) S: Select and use appropriate resources to complete the homework assignments

Developmental characteristics and grade levels	Self-regulatory processes	Applications of the role of the teacher (T) and students (S)
Grades 6–8 Increasing ability to: Plan future actions Use self-motivational strategies Engage in abstract thinking Identify a range of learning strategies	Sustain self-efficacy beliefs to master valuable tasks	Observation: T: Involves peer models in demonstration of tasks S: Observe models with similar attributes accomplish challenging tasks Emulation: T: Provides feedback to help students identify specific strategies S: Practice strategies observed by the teacher and peers Self-control: T: Provides minimal assistance while monitoring students' strategies to perform tasks S: Incorporate suggestions from teacher and peers on strategy use Self-regulation: T: Assigns independent practice that requires students to evaluate self-efficacy beliefs S: Become more self-efficacious as students are able to identify strategy use for tasks
Grades 9–12 Sustaining ability to: Master learning strategies Control social environment Delay gratification Regulate learning independently	Engage in self-evaluation of performance and outcomes	Observation: T: Demonstrates self-evaluation strategies use for performance and outcomes S: Attend to how the model uses self-evaluation strategies for performance and outcomes Emulation: T: Encourages students to discuss critical components of self-evaluation S: Describe and demonstrate the use of self-evaluation strategies as teacher scaffolds Self-control: T: Provides feedback and guidance while monitoring students' use of self-evaluation strategies S: Practice using self-evaluation strategies using teacher feedback and guidance Self-regulation: T: Encourages use of self-evaluation for tasks assigned to be completed independently S: Consistently use self-evaluation strategies to assist in independent self-evaluation

to examine the consequences of choosing to attend a party the night before a large homework assignment is due versus staying home to complete the homework.

Another educational implication is that both teachers and students need to learn and adopt appropriate help-seeking strategies. Help seeking is primarily a social process (White & Bembenutty, 2013; Karabenick & Berger, 2013; White, 2011). For instance, there are often cognitive, social, affective-emotional, and contextual resources available to teachers and students (Karabenick & Berger, 2013; Karabenick & Dembo, 2011). To illustrate, teachers can model help-seeking strategies by modeling how to determine whether there is a problem and whether help is needed, deciding whether to seek help up to the point where help is obtained and processed (Karabenick & Berger, 2013; Karabenick & Dembo, 2011).

As an initial step toward development of interventions geared at bolstering students' self-regulatory competencies, teachers should be able to first assess students' self-regulatory abilities at each of the four levels outlined in Fig. 9.2 (for reviews of ways to assess and measure self-regulated learning abilities see Boekaerts & Corno, 2005; Winne & Perry, 2000). Educators should also understand the importance and need for microanalytic assessments of self-regulated learning rather than just focusing on macroanalytic assessments. *Microanalysis assessment* is conceptualized as an umbrella term referring to highly specific or fine-grained forms of measurement targeting behaviors, cognition, or affective processes as they occur in real time across authentic contexts (Cleary, 2011a, 2011b; Cleary, Callan, & Zimmerman, 2012). Microanalytic assessments have advantages in that this methodology examines self-regulatory processes and motivation at the specific grade levels, instruction, strategy, and performance and has potential diagnostic value (Cleary, 2011b; Cleary et al., 2012; Cleary & Platten, 2013). The microanalytic methodology has also been found to be a better predictor of self-regulated learning than other previously validated measures (DiBenedetto & Zimmerman, 2013).

Finally, both teachers and students are encouraged to develop competence with computer-based learning environments because computer learning can facilitate learning, teaching, and diagnose self-regulated learning (Kitsantas & Dabbagh, 2010, 2011; Kitsantas, Dabbagh, Huie, & Dass, 2013; Greene, Moos, & Azevedo, 2011). To illustrate, 9–12 grade students could learn how to use software such as Web 2.0 technologies to promote their own learning, and teachers can use computer traces to understand self-regulatory processes of their students.

9.6.2 Implications for Future Research

Future research could investigate effective pedagogical procedures, in which modeling could be regularly integrated into the school curriculum and used in instruction of the standards. To be successful, this new line of research will need to include characteristics of competent models described by Bandura (1997). Future studies could examine whether cultural variability among learners may contribute to differences in students' self-efficacy beliefs and academic performance. Examining the

dynamic nature of the four developmental levels of self-regulated learning and the movement between the different levels is an area that future research should undertake. Further investigations into self-regulated learning processes of students' with cognitive and neurological limitations are in order, which could help these learners to be self-regulated. Research integrating new advances in technology known to be effective to enhance self-regulated learning is warranted. Future studies could also examine the effects of feedback on students' self-efficacy and self-regulation.

9.7 Conclusion

This chapter sought to examine current trends of self-regulated learning as they apply developmentally to students in different grades, and what roles teachers and learners have at each of these levels. Our discussion was guided by research on social cognitive theory with particular emphasis on the role of modeling, self-efficacy, and self-regulated learning. Specifically, we discussed Zimmerman's four developmental levels of acquiring self-regulated competency with particular attention to academic contexts in K-12 grades. We offered educational implications that we hope will help explain how learning can be enhanced in classroom settings by describing the role of teachers and students at each specific developmental level of self-regulation, and we suggested future classroom interventions for promoting self-regulated learning and future research.

Our discussion revealed that although important advances have taken place in K-12 learning settings in relation to self-regulated learning, understanding practical applications of the different developmental levels of self-regulation is still needed by both teachers and students. As our discussion uncovered, the strength of self-regulated learning is in its multi-domain scope, reaching both social and developmental domains and applicable to both educators and students across various settings. Our review suggests the important role that self-regulated learning has in the development of psychosocial skills in school systems.

References

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NY: Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.

Bandura, A. (2007). Reflections on an agentic theory of human agency. *Tidsskrift for Norsk Psykologforening, 10*, 995–1004.

Bandura, A. (2012). Social cognitive theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (Vol. 1, pp. 349–373). Thousand Oaks, CA: Sage Publications.

Bandura, A., & Mischel, W. (1965). Modifications of self-imposed delay of reward through exposure to live and symbolic models. *Journal of Personality and Social Psychology, 2*, 698–705.

Bembenutty, H. (2006, April). *Preservice teachers' help-seeking tendencies and self-regulation of learning*. A paper presented during the annual meeting of the American Educational Research Association, San Francisco, CA

Bembenutty, H. (Ed.). (2008). Self-regulated learning [Special issue]. *Journal of Advanced Academics, 20*(1)

Bembenutty, H. (Ed.). (2011). Self-regulated learning [Special issue]. *New Directions for Teaching and Learning,* (126)

Bembenutty, H. (2013). The triumph of homework completion through a learning academy of self-regulation. In H. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman*. Charlotte, NC: Information Age Publishing.

Bembenutty, H., Cleary, T. J., & Kitsantas, A. (2013). *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman*. Charlotte, NC: Information Age Publishing.

Bembenutty, H., & Karabenick, S. A. (2004). Inherent association between academic delay of gratification, future time perspective, and self-regulated learning. *Educational Psychology Review, 16*(1), 35–57.

Bembenutty, H., & White, M. C. (2013). Academic performance and satisfaction with homework completion among college students. *Learning and Individual Differences, 24*, 83–88.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review, 54*(2), 199–231.

Boekaerts, M., & Niemivirta, M. (2000). Self-regulated learning: Finding a balance between learning goals and ego-protective goals. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 417–451). San Diego, CA: Academic.

Boekaerts, M., Pintrich, P. R., & Zeidner, M. (Eds.). (2000). *Handbook of self-regulation*. San Diego, CA: Academic.

Cleary, T. J. (2011a). Professional development needs and practices among educators and school psychologists. *New Directions for Teaching and Learning, 126*, 77–87.

Cleary, T. J. (2011b). Emergence of self-regulated learning microanalysis: Historical overview, essential features, and implications for research and practice. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 329–345). New York: Routledge.

Cleary, T. J., Callan, G., & Zimmerman, B. J. (2012). Assessing self-regulation as a cyclical, context-specific phenomenon: Overview and analysis of SRL microanalytic protocols. *Education Research International, 2012*(4), 1–19.

Cleary, T. J., & Platten, P. (2013). Examining the correspondence between self-regulated learning and academic achievement: A case study analysis. *Educational Research International, 2013*(1), 1–18.

Cooper, J. E., Horn, S., & Strahan, D. B. (2005). If only they would do their homework: Promoting self-regulation in high school English classes. *High School Journal, 88*(30), 10–25.

Dembo, M. H. (2001). Learning to teach is not enough: Future teachers also need to learn how to learn. *Teacher Education Quarterly, 28*(4), 23–35.

DiBenedetto, M. K., & Bembenutty, H. (2011). Differences between full-time and part-time MBA students' self-efficacy for learning and employment: A self-regulatory perspective. *The International Journal of Educational and Psychological Assessment, 7*(1), 81–110.

DiBenedetto, M. K., & White, M. C. (2013). Applying the model of development of self-regulatory competence to mentoring. In H. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman* (pp. 445–472). Charlotte, NC: Information Age Publishing.

DiBenedetto, M. K., & Zimmerman, B. J. (2010). Differences in self-regulatory processes among students studying science: A microanalytic investigation. *The International Journal of Educational and Psychological Assessment, 5*(1), 2–24.

DiBenedetto, M. K., & Zimmerman, B. J. (2013). Construct and predictive validity of microanalytic measures of students' self-regulation of science learning. *Learning and Individual Differences, 26*, 30–41.

Dignath, C., Buettner, G., & Langfeldt, H. (2008). How can primary school students learn self-regulated learning strategies most effectively? A meta-analysis on self-regulation training programmes. *Educational Research Review, 3*(2), 101–129.

Dorn, L. J. & Soffos, C. (2001). *Shaping literate minds: Developing self-regulated learners*. Portland, ME: Stenhouse Publishers.

Eccles, J. S., Midgley, C., & Adler, T. F. (1984). Grade-related changes in the school environment: Effects on achievement motivation. In J. Nicholls (Ed.), *Advances in motivation and achievement* (Vol. 3, pp. 283–331). Greenwich, CT: JAI Press.

Garcia, T., & Pintrich, P. R. (1996). The effects of autonomy on motivation and performance in the college classroom. *Contemporary Educational Psychology, 21*(4), 477–486.

Graham, S., & Harris, K. R. (1989). Improving learning disabled students' skills at composing essays: Self-instructional strategy training. *Exceptional Children, 56*, 201–216.

Greene, J. A., Moos, D. C., & Azevedo, R. (2011). Self-regulation of learning with computer-based learning environments. *New Directions for Teaching and Learning, 126*, 107–115. doi:10.1002/tl.449.

Horner, S. L., & Shwery, C. S. (2002). Becoming an engaged, self-regulated reader. *Theory Into Practice, 41*(2), 102–109.

Karabenick, S. A., & Berger, J. (2013). Help seeking as a self-regulated learning strategy. In H. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman* (pp. 237–261). Charlotte, NC: Information Age Publishing.

Karabenick, S. A., & Dembo, M. H. (2011). Understanding and facilitating self-regulated help seeking. *New Directions for Teaching and Learning, 126*, 33–44.

Kitsantas, A., & Dabbagh, N. (2010). *Learning to learn with integrative learning technologies (ILT): A practical guide for academic success*. Greenwich, CT: Information Age Publishing.

Kitsantas, A., & Dabbagh, N. (2011). The role of Web 2.0 technologies in self-regulated learning. *New Directions for Teaching and Learning, 126*, 99–106.

Kitsantas, A., Dabbagh, N., Huie, F., & Dass, S. (2013). Learning technologies and self-regulated learning: Implications for practice. In H. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman* (pp. 325–354). Charlotte, NC: Information Age Publishing.

Kitsantas, A., Zimmerman, B. J., & Cleary, T. J. (2000). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology, 94*, 660–668.

Lipnevich, A. A., MacCann, C., Krumm, S., Burrus, J., & Roberts, R. D. (2011). Mathematics attitudes and mathematics outcomes of U.S. and Belarusian middle school students. *Journal of Educational Psychology, 103*(1), 105–118.

Lipnevich, A. A., & Smith, J. K. (2009). Effects of differential feedback on students' examination performance. *Journal of Experimental Psychology: Applied, 15*(4), 319–333.

Lombaerts, K., Engels, N., & van Braak, J. (2008). Determinants of teachers' recognitions of self-regulated learning practices in elementary education. *Journal of Educational Research, 103*(3), 163–173.

McPherson, G. E., & Davidson, J. W. (2002). Musical practice: Mother and child interactions during the first year of learning an instrument. *Music Education Research, 4*(1), 141–156.

Nedwman, R. S. (2000). Social Influences on the Development of Children's Adaptive Help Seeking: The Role of Parents, Teachers, and Peers. *Developmental Review, 20*(3), 350–404. http://dx.doi.org/10.1006/drev.1999.0502

Ormrod, J. E. (2011). Essentials of Educational Psychology: Big Ideas to Guide Effective Teaching (3rd ed.). New York, NY: Pearson.

Ormrod, J. E. (2013). Essentials of Educational Psychology: Big Ideas to Guide Effective Teaching (8th ed.). New York, NY: Pearson.

Pajares, F. (1996). Self-efficacy beliefs in achievement settings. *Review of Educational Research, 66*, 543–578.

Pape, S. J., Bell, C. V., & Yetkin-Ozdemir, I. E. (2013). Sequencing components of mathematics lessons to maximize development of self-regulation: Theory, practice, and intervention.

In H. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman* (pp. 29–58). Charlotte, NC: Information Age Publishing.

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist, 36*, 89–101.

Perry, N., & VandeKamp, K. (2000). Creating classroom contexts that support young children's development of self-regulated learning. *International Journal of Educational Research, 33*, 821–843.

Perry, N. E., VandeKamp, K. O., Mercer, L. K., & Nordby, C. J. (2002). Investigating Teacher-Student Interactions That Foster Self-Regulated Learning. *Educational Psychologist, 37*(1), 5-15. http://dx.doi.org/10.1207/S15326985EP3701_2

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M., Boekaerts, & P. R., Pintrich. (Eds.) *Handbook of Self-regulation* (pp.451–502). San Diego: Academic Press.

Puustinen, M., & Pulkkinen, L. (2001). Models of self-regulated learning: A review. *Scandinavian Journal of Educational Research, 45*(3), 269–286.

Ramdass, D. (2011). Enhancing mathematics skill and self-regulatory competency through observation and emulation. *The International Journal of Research and Review, 7*(1), 24–44.

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research, 57*, 149–174.

Schunk, D. H., & DiBenedetto, M. K. (2014). Academic self-efficacy. In M. J. Furlong, R. Gillman, & E. S. Huebner (Eds.), *Handbook of positive psychology in the schools* (2nd ed., pp. 115–130). New York: Routledge.

Schunk, D. H., & DiBenedetto, M. K. (2015). Self-efficacy: Educational aspects. In J. D. Wright (Ed.), *International encyclopedia of social and behavioral sciences* (2nd ed.). Oxford, UK: Elsevier.

Schunk, D. H., & Ertmer, P. A. (2000). Self-regulation and academic learning: Self-efficacy enhancing interventions. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 631–649). San Diego, CA: Academic.

Schunk, D. H., & Mullen, C. A. (2013). Self-efficacy as an engaged learner. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 219–235). New York: Springer.

Schunk, D. H., & Pajares, F. (2009). Self-efficacy theory. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation in schools* (pp. 35–53). New York: Routledge.

Schunk, D. H., & Rice, J. M. (1993). Strategy fading and progress feedback: Effects on self-efficacy and comprehension among students receiving remedial reading services. *Journal of Special Education, 27*, 257–276.

Schunk, D. H., & Usher, E. L. (2012). Social cognitive theory and motivation. In R. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 13–27). New York: Oxford University Press.

Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist, 3*(4), 195–208.

Troia, G. A., & Graham, S. (2002). The effectiveness of a highly explicit, teacher-directed strategy instruction routine. *Journal of Learning Disabilities, 35*(4), 290–306.

Usher, E. L., & Pajares, F. (2006). Sources of academic and self-regulatory efficacy beliefs of entering middle school students. *Contemporary Educational Psychology, 31*, 125–141.

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*, 751–796.

Vohs, K. D. & Baumeister, R. F. (2011). *Handbook of Self-Regulation: Research, Theory, and Applications* (2nd edition). New York, NY: Guilford.

White, M. C. (2011). Predicting success in teacher certification testing: The role of academic help-seeking. *The International Journal of Educational and Psychological Assessment, 7*, 22–44.

White, M. C., & Bembenutty, H. (2013). Not all avoidance help seekers are created equal individual differences in adaptive and executive help seeking. *SAGE Open, 3*(2), 1–14.

White, M. C., & DiBenedetto, M. K. (2015). *Self-regulation and the common core: Application to ELA standards*. New York: Routledge.

Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement* (pp. 153–190). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Winne, P. H. (2011). A cognitive and metacognitive analysis of self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation and performance* (pp. 15–32). New York: Taylor & Francis.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 532–566). San Diego, CA: Academic.

Zimmerman, B. J. (2000). Attainment of self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego, CA: Academic.

Zimmerman, B. J. & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology, 94*(4), 660-668. http://dx.doi.org/10.1037/0022-0663.94.4.660

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist, 48*, 1–13.

Zimmerman, B. J., Bonner, S., & Kovach, R. (1996). *Developing self-regulated learners: Beyond achievement to self-efficacy*. Washington, DC: American Psychological Association.

Zimmerman, B. J., & DiBenedetto, M. K. (2008). Mastery learning and assessment: Students and teachers in an era of high stakes testing. *Psychology in the Schools, 45*(3), 206–216.

Zimmerman, B. J., & Kitsantas, A. (1997). Developmental phases in self-regulation: Shifting from process goals to self-regulatory goals. *Journal of Educational Psychology, 91*, 241–250.

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2011). *Handbook of self-regulation of learning and performance*. New York: Routledge.

Chapter 10
Student Motivation: Current Theories, Constructs, and Interventions Within an Expectancy-Value Framework

Chris S. Hulleman, Kenneth E. Barron, Jeff J. Kosovich, and Rory A. Lazowski

10.1 Introduction

Amanda and Rachel are two students who, on the surface, look quite similar as they begin their first year of high school. They attend the same school and were placed into all of the same classes with the same teachers. However, by the time they graduate, they will look very different. Amanda will barely earn a high school diploma, whereas Rachel will graduate with honors and have her pick of which university to attend to continue her education. The question is why. Why does Amanda struggle and Rachel thrive? Importantly, what could teachers and the school have done to intervene to change Amanda's outcome?

The answer lies in how Amanda and Rachel respond to academic challenges. In response to lagging performance on international tests, as well as low graduation rates, K-12 school systems have been pushed to increase academic expectations and standards. However, without proper support for students, increasing standards for student learning heightens the risk that more students will fail and leave the educational system (Dweck, Walton, & Cohen, 2011). This raises an important

C.S. Hulleman (✉) • J.J. Kosovich
Department of Educational Leadership, Policy, and Foundations, University of Virginia, Charlottesville, VA, PO Box 400877, USA
e-mail: chris.hulleman@virginia.edu; jjkosy@gmail.com

K.E. Barron
Department of Psychology, James Madison University, MSC 7704, 800 S. Main St., Harrisonburg, VA, USA
e-mail: barronke@jmu.edu

R.A. Lazowski
K-12 Research, The College Board, 800 Township Line Road, Suite 200, Yardley, PA 19067, USA
e-mail: rlazowski@collegeboard.org

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_10

question: Even though adults see these standards and associated learning activities as necessary and important to facilitate student learning, why would students be motivated to engage in these academic tasks? After all, having rigorous standards means that more pressure is placed on students to exceed expectations on increasingly difficult academic tasks. So understanding how students approach more rigorous tasks, and their accompanying type and quantity of motivation, is an essential component of school reform efforts. That is, without attending to student perspectives, how will increases in rigor lead to desired learning outcomes, such as increased performance on standardized tests, and the kind of deep learning that enables complex problem solving to occur? Further, when students fail to respond to increased challenge and pressure, what can educators do to increase student engagement in learning?

The motivation for students to enter into a setting where achievement is highly valued, and the degree to which students engage in the associated tasks and activities, is known as achievement motivation (Weiner, 1980). Understanding the development of achievement motivation, why individuals differ in achievement motivation, the outcomes associated with higher and lower levels of achievement motivation, and what contextual factors amplify or impede achievement motivation is a central task for both researchers and practitioners.

Within achievement motivation research, an expectancy-value framework has been particularly generative. From their earliest psychological roots, expectancy-value models focused on understanding the factors that predicted behavior within situations where individuals were trying to achieve an outcome (Atkinson, 1958; Lewin, Dembo, Festinger, & Sears, 1944). However, these initial approaches were context-free; that is, much of the research and thinking was focused on arbitrary tasks in laboratory settings (Weiner, 1980). It wasn't until the early 1980s that Jacqueline Eccles and her colleagues proposed a modern version of expectancy-value motivation focused on students' achievement choices within educational contexts (Eccles et al., 1983).

Not only did the Eccles framework bring expectancy and value constructs to prominence in explaining achievement behavior in educational contexts, it also highlighted two critical aspects of motivation that are necessary for students to be optimally engaged. First, students need to believe that they can succeed (i.e., they need to have positive expectancies). Second, students need to perceive an important reason to engage in the behavior (i.e., they need to have positive values). For example, in Amanda's struggle in school, is it an expectancy issue, where Amanda begins to doubt her ability to successfully complete her school work? Is it a value issue, where she fails to see a reason or purpose for her coursework? Or is it some combination of both?

Although the expectancy-value framework offers a multidimensional approach to understanding student motivation, both expectancy and value have their own rich bodies of literature. In fact, one of the original motivators of Eccles and her colleagues (Eccles et al., 1983; Parsons et al., 1980) was to adopt a theoretical model that integrated findings from multiple theoretical perspectives. Thus, in an effort to help organize understanding of this research area, the first purpose of our chapter is

to offer separate reviews of expectancy and value constructs. The second purpose of our chapter is to consider integrative approaches that combine expectancy and value constructs within the same model. The final purpose of our chapter is to highlight an emerging body of intervention work designed to enhance students' expectancies and values. By identifying the sources of expectancy and value amenable to change, we can help practitioners diagnose why students like Amanda struggle and how teachers and schools can purposefully increase student motivation.

10.2 Review of Theoretical Constructs and Research on Expectancy-Related Beliefs

Theories concerned with expectancy-related constructs attempt to address the first critical question about motivation: "Do students think they can do the task?" There are numerous theoretical conceptualizations (e.g., self-efficacy theory, expectancy-value theory, locus of control theories, attribution theories, and implicit theories of intelligence) and specific constructs (e.g., self-efficacy beliefs, expectancies for success, perceived control, perceptions of task difficulty, and growth mindsets) linked to addressing this motivational question (see Pajares, 1996; Pintrich, 2003). Although there are similarities among these theories and proposed constructs, there are also substantive, theoretical differences that distinguish each. We review different theories and constructs briefly below and offer an overall summary in Table 10.1.

10.2.1 Self-efficacy

Bandura (1997) conceptualized *self-efficacy* as a belief in one's ability to plan and execute the skills necessary to produce certain behaviors. Bandura also distinguished *self-efficacy beliefs* from *outcome expectancies*. Whereas self-efficacy beliefs are related to whether an individual can successfully complete a task (e.g., learn how to solve a particular math problem), outcome expectancies are related to whether an individual can successfully obtain a particular outcome or consequence of accomplishing the task (e.g., get an A on a math test). Self-efficacy beliefs are proposed to be determined by previous performance, vicarious learning (observing a model successfully complete a task), verbal encouragement from others, and physiological reactions to a situation or task (Bandura, 1997). Previous performance of a skill is considered a strong source of self-efficacy, representing tangible, authentic evidence that an individual can or cannot perform the requisite skill. Vicarious learning (i.e., seeing others perform a task successfully) also increases self-efficacy, and observing models closer to the individual's peer group is typically more effective than observing an expert performing the skill. When an individual receives positive verbal encouragement from a knowledgeable and reliable source (such as a teacher), then self-efficacy tends to increase. Finally, more positive physiological reactions, such

Table 10.1 Expectancy-related constructs and measures

Construct	Sample item
1. Self-efficacy	
Self-efficacy beliefs[a]	Rate the probability of successfully performing the following task from zero (no chance) to 100 (complete certainty): Writing a one- or two-sentence answer to a specific test question.
Outcome expectancies[a]	How important is writing for getting good grades in school?
Content specific examples:	
Self-efficacy for self-regulated learning[b]	How well can you finish homework assignments by deadlines?
Statistics self-efficacy[c]	How confident are you that you can identify the factors that influence power?
Teaching efficacy[d]	How much can you influence the decisions that are made in your school?
2. Expectancies	
Ability beliefs[e]	How good at reading/English are you?
Expectancies for success[e]	How well do you expect to do in reading/English next year?
3. Self-concept	
Math self-concept[f]	Mathematics makes me feel inadequate (reverse scored).
Self-concept of ability[g]	How well do you expect to do in (domain X) this year?
General school self-concept[h]	I learn things quickly in most school subjects.
4. Perceived control	
External locus of control[i]	Many times exam questions tend to be so unrelated to course work that studying is really useless.
Internal locus of control[i]	In the case of the well-prepared student, there is rarely if ever such a thing as an unfair test.
Strategy beliefs – effort[j]	If I want to do well on my schoolwork, I need to try hard.
Strategy beliefs – ability[j]	If I am not smart, I won't get good grades.
Perceived control – situational level[k]	I have a great deal of control over my academic performance in my psychology class.
5. Attributions	
Causality ability[l]	If I were to receive low marks, it would cause me to question my academic abilities.
Effort[l]	In my case, the good grades I receive are always the direct result of my efforts.
Context[l]	Often my poorer grades are obtained in courses that the professor has failed to make interesting.
Luck[l]	Sometimes my success on exams depends on some luck.
6. Implicit theories of intelligence	
Incremental theory[a]	You can always greatly change how intelligent you are.
Entity theory[m]	You have a certain amount of intelligence, and you really can't do much to change it.

[a]Shell, Murphy, and Bruning (1989); [b]Bandura (1989); [c]Finney and Schraw (2003); [d]Bandura (1993); [e]Durik, Vida, and Eccles (2006); [f]Marsh (1992); [g]Denissen, Zarrett, and Eccles (2007); [h]Marsh (1990); [i]Rotter (1966); [j]Patrick, Skinner, and Connell (1993); [k]Stupnisky, Perry, Hall, and Guay (2012); [l]Lefcourt, von Baeyer, Ware, and Cox (1979); [m]Dweck (1999)

as feelings of excitement, can accompany higher levels of self-efficacy, whereas anxious reactions can accompany lower levels of self-efficacy (for a review see Usher & Pajares, 2008; see also Chaps. 9 and 11).

A large body of research has linked self-efficacy with educational outcomes (Haney & Durlak, 1998). Although a complete description of these associations is beyond the scope of this chapter, examples include positive associations with goal setting (Locke & Latham, 1990), self-regulation (Chap. 9; Zimmerman, Bandura, & Martinez-Pons, 1992), and effort, persistence, and resilience (Pajares, 2002; Robbinset al., 2004; Schunk & Pajares, 2002). In particular, self-efficacy predicts educational outcomes most closely aligned with the referent task. For example, Bong and Skaalvik (2003) found that math self-efficacy was the only predictor of math performance and English self-efficacy was the only predictor of English performance, even when these and other achievement indexes were included in the model (cf. Baranik, Barron, & Finney, 2010). As a result, self-efficacy is often measured at a task- or subject-specific level (see Table 10.1).

10.2.2 Expectancies

Eccles and colleagues (1983) defined expectancies as an individual's perceptions about whether he or she can successfully accomplish a task. In their model, they also proposed two types of expectancies: *ability beliefs* and *expectancies for success* (Eccles & Wigfield, 2002). Ability beliefs refer to a person's current sense of competence in being able to complete a task. In contrast, expectancies for success reflect how successful an individual believes he or she can continue to be in the future. Although these two types of expectancies are theoretically distinguishable (reflecting separate beliefs about one's current ability and future performance, respectively), empirical attempts to measure them separately have not been supported, resulting in one, overall expectancy scale (Wigfield & Eccles, 2000).

Although similar to the construct of self-efficacy, there are important theoretical distinctions that can be drawn between expectancies and self-efficacy (Pajares, 1996). Expectancies relate to more general or broad domains and in turn more strongly relate to general or broad outcomes. Self-efficacy, on the other hand, focuses on more specific tasks that correspond to being able to achieve a specific result. For example, an expectancy measure may evaluate an individual's capabilities in a certain subject area (e.g., English), and this measure may be used to predict course grades in that subject. In contrast, a self-efficacy measure may evaluate an individual's capabilities to perform a specific task within a class (e.g., being able to appropriately use commas when writing), and the responses may be used to predict actual performance on this specific task.

In general, research suggests that expectancies most strongly predict student achievement, such as test scores, course grades, and GPA (e.g., Eccles et al., 1983; Richardson, Abraham, & Bond, 2012; Robbins et al., 2004; Xiang, Chen, & Bruene, 2005), but also predict choice of course enrollment, persistence, career aspirations, and task engagement (e.g., Durik, Vida, & Eccles, 2006; Robbins et al., 2004).

10.2.3 Self-concept

Self-concept, broadly defined, is an individual's perception of themselves (Shavelson, Hubner, & Stanton, 1976; Chap. 8) and is largely regarded as a multidimensional construct that includes academic and nonacademic forms of self-concept (Marsh, 1990). Together, these more specific types of self-concept form a collective self-concept in the individual. As such, self-concept is considered hierarchical, with a general self-concept formed by both academic and nonacademic components (Marsh, 1990). For the academic domain, research suggests that verbal and math self-concepts might not necessarily form a single dimension together (Byrne & Shavelson, 1986). Therefore, current conceptions specify separate verbal and math self-concepts in addition to a domain general academic self-concept (Marsh & Shavelson, 1985).

Positive academic self-concept has been associated with higher levels of achievement, particularly for grades and standardized test scores, while controlling for previous achievement (Marsh, Byrne, & Yeung, 1999). Academic self-concept also has been shown to predict teacher ratings of student engagement and persistence (Skaalvik & Rankin, 1996; Skinner, Wellborn, & Connell, 1990), self-reported effort (Skaalvik & Rankin, 1995), and adaptive help-seeking behaviors (Ames, 1983). Furthermore, Marsh and Martin (2011) found that the relationships are reciprocal. For instance, positive educational outcomes can enhance self-concept just as higher levels of self-concept can yield more positive outcomes.

Once again, the distinctions between self-concept and other expectancy-related constructs correspond to levels of specificity. Self-concept at the more general level is similar to the more broad ability beliefs, whereas the more specific types of self-concepts are more aligned with self-efficacy and tend to be more distinct. The level of specificity corresponds to the predictive power of self-concept, with more specificity in the self-concept construct and outcome yielding more powerful results. In two separate meta-analyses of academic self-concept and academic achievement, Huang (2011) and Richardson et al. (2012) found that the effect of self-concept was smaller on achievement when studies used more global, as compared to subject-specific, measures of academic self-concept.

10.2.4 Perceived Control

Rotter (1966) first proposed the term *locus of control* to describe the perceived control an individual possesses over certain outcomes. An individual's locus of control can comprise one of two types: *external locus of control* and *internal locus of control*. If an individual perceives that outcomes occur due to factors outside of his or her control (such as luck or fate), then that individual maintains an *external locus of control*. In contrast, if an individual perceives that outcomes occur due to factors within his or her control or capacity (such as effort), then that individual holds an

internal locus of control. Locus of control is hypothesized to be a continuum between internal and external forms and is influenced by environmental, cultural, and personal variables (Rotter, 1966).

Skinner (1996) proposed that individuals develop their locus of control over time with repeated behavior-outcome contingencies. This might include how individuals perceive that certain behaviors subsequently lead to favorable or unfavorable outcomes. These associations thus inform the individual as to the level of control they have over future outcomes. Over time, the habitual endorsements individuals ascribe to these outcomes reinforce attributions and their locus of control (Weiner, 2010). Research supports that higher levels of an internal locus of control are associated with higher levels of academic achievement (Skinner, 1995), perceptions of competence (Connell & Wellborn, 1991), and hours spent studying (Bodill & Roberts, 2013), particularly for students whose perceptions remain stable over time (Stupinsky, Perry, Hall, & Guay, 2012). In contrast, higher levels of an external locus of control are associated with higher levels of anxiety, less autonomy, lower levels of motivation to make behavioral changes (Lavender, 2005), and fewer hours spent studying (Bodill & Roberts, 2013).

In addition to the internal/external distinction, other researchers have further differentiated control beliefs. For example, Skinner and colleagues (Skinner, Wellborn, & Connell, 1990) proposed three types of perceived control beliefs: (1) *means-end* or *strategy beliefs* ("the extent that potential causes produce given outcomes"; Schunk, 1991, p. 208), (2) *agency* or *capacity beliefs* ("whether the individual has or can acquire the potential causes"; Schunk, 1991, p. 208), and (3) *control beliefs* ("whether the individual can produce the desired outcome without reference to any particular means"; Schunk, 1991, p. 208). In the control-value model, Pekrun (2006; Goetz & Bieg, 2016) proposed two dimensions of control: attributions for past success/failure and expectations for future success. Research from both of these perspectives reveals positive associations between control and achievement outcomes, including performance and emotions.

10.2.5 Attributions

Weiner (1972) proposed that the attributions individuals ascribe to success or failure have particular bearing on expectancies and associated educational outcomes. For example, attribution theory posits that individuals frequently attribute success and failure to perceived causes such as ability, effort, perceived task difficulty, or luck. Like theories of perceived control, ability and effort are considered to lie within the individual (*internal*), whereas perceived task difficulty and luck are considered to lie outside of the individual (*external*). However, in addition to an internal vs. external locus of causality dimension, Weiner (2010) further differentiates attributions into stable vs. unstable and controllable vs. uncontrollable dimensions. Perceived causes like ability and task difficulty are consistent across contexts (*stable*), whereas effort and luck are more variable across contexts and potentially unpredictable (*unstable*).

Moreover, effort and task difficulty can be influenced directly by the student and teacher (*controllable*), whereas current ability and luck cannot (*uncontrollable*).

Expectancies for success will increase if the individual perceives that successful completion of a task is a result of causal factors that are internal and stable – that is, the conditions that resulted in the successful outcome are likely to occur again in the future and are within his or her control. Conversely, expectancies for success will not increase if successful task completion is attributed to external and unstable factors because of the instability surrounding the conditions that caused the outcome to occur and the belief that these factors are outside of one's control (see Weiner, Nierenberg, & Goldstein, 1976). Weiner (2010) also stressed the importance of the associations between attributions and emotions as they relate to success or failure depending on internal/external, stable/unstable, and controllable/uncontrollable dimensions. These include feelings such as pride (internal-success), guilt/regret (internal-controllable-failure), shame/humiliation (internal-uncontrollable-failure), hopelessness (stable-failure), and hope (unstable-failure). These hypotheses are generally supported by the research literature which reveals that attributing success to external, unstable causes – compared to internal, controllable causes – is associated with worse achievement outcomes (e.g., Glasgow, Dornbusch, Trover, Steinberg, & Ritter, 1997) and emotional well-being (Ciarrochi, Heaven, & Davies, 2007).

10.2.6 Implicit Theories of Intelligence

Implicit theories of intelligence (see Dweck & Leggett, 1988; Dweck, 1999) posit that individuals generally possess one of two different theories regarding their intelligence: (1) that intelligence is dynamic, malleable, and amenable to change given sufficient effort and hard work (*incremental view of intelligence*) or (2) that intelligence is fixed, static, and resistant to change regardless of effort and hard work (*entity view of intelligence*). In more recent writing, Dweck (2006) refers to these two views as a growth or fixed mindset, respectively. These mindsets may develop as a result of the messages individuals receive from parents and teachers (Mueller & Dweck, 1998) and can therefore be manipulated or changed (e.g., Blackwell, Trzesniewski, & Dweck, 2007; Good, Aronson, & Inzlicht, 2003). Thus, environmental factors and social structures are important sources contributing to an individual's theory of intelligence.

Whether or not individuals attribute their success or failure on a task to effort (incremental view) or fixed ability (entity view) substantially impacts a number of outcomes. For instance, correlational field studies have demonstrated that those with higher levels of incremental views of intelligence are more likely to focus on mastery/learning goals (Dweck & Leggett, 1988), have higher levels of task persistence and task enjoyment (Mueller & Dweck, 1998), and have better long-term academic performance (Aronson, Fried, & Good, 2002; Blackwell et al., 2007; Good et al., 2003; Romero et al., 2014). Furthermore, a recent meta-analysis conducted by Burnette, O'Boyle, VanEpps, Pollack, and Finkel (2013) found that

incremental views of intelligence predicted higher levels of mastery goals, use of more mastery-oriented self-regulation strategies (and less use of helpless-oriented strategies), and higher levels of effective goal monitoring.

The particular theory of intelligence that an individual endorses has bearing on constructs discussed earlier. For example, Dweck and Leggett (1988) noted that entity and incremental theories represent different forms of self-concept. With an entity view, self-concept would be regarded as a collection of fixed characteristics that could reliably be measured and evaluated. With an incremental view, however, self-concept would be regarded as a collection of changeable characteristics that would change over time as a result of concerted effort.

10.2.7 Expectancy-Related Beliefs and Learning Outcomes

Expectancy-related constructs have historically shown strong associations with effort, persistence, achievement, and engagement (Pintrich, 2003). Additionally, many of the expectancy constructs demonstrate a reciprocal relationship with these outcomes, such that expectancy beliefs affect certain outcomes and these outcomes in turn affect subsequent expectancy (e.g., Bandura, 1997; Eccles & Wigfield, 2002; Marsh & Martin, 2011).

Although the relationship between expectancies and outcomes is hypothesized to be present across grade level, the magnitude of the relationship may vary (for reviews see Eccles & Wigfield, 2002; Wigfield & Eccles, 1992, 2000). Unfortunately, there have been no systematic reviews of the relationship between expectancies and outcomes across student grade level. For example, Usher and Pajares (2008) systematically reviewed the self-efficacy literature and did not make conclusions based on age differences because the majority of studies focused on high school and college students. Future research syntheses need to systematically examine age-related differences in the relationship between expectancies and outcomes, and future research needs to explicitly examine age as a moderating factor of expectancy-outcome relationships.

Instead, there have been several meta-analytic reviews that have included expectancy-related constructs at the college level, and their results are worth mentioning here. For example, in their meta-analysis of 13 years of research on the antecedents of college students' GPA, Richardson and colleagues (2012) found that measures of performance expectations and academic self-efficacy/self-concept had the strongest correlations with GPA among all the psychosocial factors included in the paper. In addition, self-efficacy accounted for unique variance in GPA when controlling for high school GPA and SAT/ACT. A separate meta-analysis by Robbins and colleagues (2004) mirrored these results on GPA. In addition, the authors also examined predictors of first-year retention in college and found that academic self-efficacy had the strongest correlation among all psychosocial predictors, accounting for unique variance in persistence when controlling for high school GPA, ACT/SAT, and socioeconomic status.

10.3 Review of Theoretical Constructs and Research on Values

Whereas expectancy-related beliefs focus on the first critical question of motivation, "Can students do the task?," value-related constructs address the second critical motivational question: "Do students want to do the task?" Once again, numerous theoretical conceptualizations (e.g., expectancy-value theory, self-determination theory, and self-worth theory) and specific constructs (e.g., intrinsic value, attainment value, utility value, and extrinsic value) have been proposed (see Pintrich, 2003). We review each of these theories and constructs briefly below and offer an overall summary in Table 10.2.

10.3.1 Subjective Task Values

Subjective task values, as defined by expectancy-value theory, are considered one of the most proximal determinants of achievement behavior. Eccles and colleagues (1983) and Eccles (2005) have consistently proposed four major types: *intrinsic value*, *utility value*, *attainment value*, and *cost*. Together, these different types of task value combine to guide task engagement. In the Eccles and colleagues' model, task values are considered to be subjective because the value of a task is dictated by how an individual perceives and appraises the task.

Intrinsic Value Intrinsic value (also called interest value) is defined as the inherent enjoyment or satisfaction an individual perceives that he or she will obtain from engaging in a task (Eccles et al., 1983). Perceiving an academic task as being intrinsically valuable and interesting leads to focused attention, deeper information processing, and increased learning outcomes (Renninger & Hidi, 2011).

Intrinsic value has been related to a number of other motivational factors. For example, it was found that people who initially set goals to learn material (rather than simply demonstrate competence) reported higher subsequent intrinsic value, task satisfaction, and interest (Hulleman, Durik, Schweigert, & Harackiewicz, 2008). In their study examining the relationship between situational interest, enjoyment, and persistence, Fulmer and Tulis (2013) found that individuals experiencing more enjoyment and situational interest (i.e., intrinsic value) reported more persistence on a task.

Utility Value Utility value is defined as the usefulness or importance of a particular task to an individual's current or future goals (Eccles et al., 1983). Gaspard et al. (2015) identified five different subtypes of utility value. One of the most prominent subcomponents, *perceived instrumentality*, is specifically oriented toward future goals and pursuits (De Volder & Lens, 1982). Other identified components included *social utility*, *utility for school* (Conley, 2012), *utility for daily life* (Hulleman & Harackiewicz, 2009), and *career utility* (Hulleman et al., 2008). The focus on personal

Table 10.2 Value constructs and measures

Construct	Sample item
1. Task values	
Intrinsic value[a]	I enjoy coming to lecture.
Utility value[b]	This technique could be useful in everyday life.
Attainment value[c]	I feel that, to me, being good at solving problems which involve science or reasoning scientifically is: 1 (not at all important) to 6 (very important).
Cost[c]	When I think about the hard work needed to get through my science major [or science track], I am not sure that getting a science degree is going to be worth it in the end.
2. Intrinsic and extrinsic motivation	
Intrinsic motivation[d]	I work really hard because I like to learn new things.
Extrinsic motivation[d]	I work on problems because I'm supposed to.
Motivation regulation continuum	
Amotivation[e]	I don't know why [I go to school]; I can't really see what good it will do for me.
External regulation[e]	[I go to school] because my parents pressure me to go.
Introjected regulation[e]	[I go to school] because if I did not go I'd be angry with myself for a long time.
Identified regulation[e]	[I go to school] because I feel that postsecondary studies will help me to prepare myself for the career I have chosen.
Integrated regulation[e]	[I go to school] because in choosing to continue to study, I'll be the type of person that will be in a better situation to get better job opportunities.
Intrinsic regulation[e]	[I go to school] because I experience pleasure and satisfaction in learning new things.
Psychological needs	
Competence[f]	During this event I felt very capable in what I did.
Autonomy[f]	During this event I felt free to do things my own way.
Relatedness[f]	During this event I felt close and connected with other people who are important with me.
Self-esteem[f]	During this event I felt quite satisfied with who I am.
Self-worth[g]	I don't care if other people have a negative opinion about me.
3. Human values	
Terminal value[h]	As the guiding principle in my life: wisdom [is 7 (of supreme importance) to 0 (opposed to my values)].
Instrumental value[h]	As the guiding principle in my life: intellect [is 7 (of supreme importance) to 0 (opposed to my values)].

[a]Hulleman, Durik, Schweigert, & Harackiewicz, (2008); [b]Hulleman, Godes, Hendricks, & Harackiewicz, (2010); [c]Perez, Cromley, & Kaplan, (2014); [d]Lepper, Henderlong, Corpus, & Iyengar, (2005); [e]Vallerand and Blssonnette (1992); [f]Sheldon, Elliot, Kim & Kasser, (2001); [g]Crocker, Luhtanen, Cooper, and Bouvrette (2003); [h]Schwartz (1994)

meaning for important life goals inherent in utility value helps separate it from intrinsic value, which is more focused on enjoyment. Thus, an academic activity can lack intrinsic value yet have high utility value. For example, a student majoring in biology may not gain much enjoyment from chemistry (i.e., intrinsic value); however, learning chemistry may be valuable because it enables the student to pursue a medical degree (i.e., utility value) or solve an important social problem, such as creating clean water in impoverished nations (Yeager & Bundick, 2009).

Whereas intrinsic value tends to be related primarily to choice-related outcomes, utility value has also been positively linked to performance outcomes (Hulleman, Godes, Hendricks, & Harackiewicz, 2010; Hulleman et al., 2008; Hulleman & Harackiewicz, 2009; Simons, DeWitte, & Lens, 2003). When students perceive a task as more instrumental to their future goals, they are more likely to be persistent and also reach higher levels of achievement (De Volder & Lens, 1982). Moreover, students who focus on intrinsic, future goals are more excited and persistent and demonstrate superior performance (Simons et al., 2003). For example, Updegraff, Eccles, Barber, and O'Brien (1996) found that utility value for mathematics predicted the number of high school math courses taken even when controlling for math GPA, aptitude, and self-concept.

Attainment Value Attainment value is often described as the importance of a task to an individual's self-concept or identity (Eccles, 2009; Eccles & Wigfield, 2002). For example, a student may report high attainment value for math class because demonstrating competence in that situation affirms one's identity as a math major. Attainment value was adapted from early value research that defined it as the importance an individual attaches to competent performance in a specific area of achievement (e.g., Crandall, Katkovsky, & Preston, 1962). When attainment value was subsumed within the expectancy-value framework, it was eventually defined so that it focused on an individual's identity (Eccles, 2009). Attainment value has been found to be positively correlated with intrinsic and utility value, cognitive engagement, and intentions to continue education (e.g., Battle & Wigfield, 2003; Eccles, Wigfield, Harold, & Blumenfeld, 1993; Johnson & Sinatra, 2013; Wigfield et al., 1997).

Cost Instead of the positive aspects of wanting to do an activity, cost reflects the perceived negative aspects of a task. These negative perceptions discourage an individual from engaging in the activity, decrease persistence while engaging in the activity, and lead to a devaluing of the activity (Eccles et al., 1983). Three major sources of cost have consistently been proposed in the literature: the amount of effort required by a task, how engaging in one task results in missing out on other valued alternatives, and the negative emotional states that occur while doing the task. Eccles et al. (1983) suggested that the first two types of cost represent the costs of success (e.g., giving up your time and energy or giving up other valued activities), and the third reflects the costs of failing (e.g., anxiety).

New theoretical and empirical work on the construct of cost suggests that it is a separate construct on par with both value and expectancy (Barron & Hulleman,

2015; Flake, Barron, Hulleman, McCoach, & Welsh, 2015; Kosovich, Hulleman, Barron, & Getty, 2015) that can directly influence an individual's overall motivation. For example, recent research has shown that cost is negatively related to achievement outcomes (e.g., Conley, 2012; Grays, 2013; Kosovich et al., 2015; Perez, Cromley, & Kaplan, 2014; Trautwein et al., 2012). This enables us to consider cost as a distinct source of motivation, in addition to expectancy and value, that could be ameliorated to benefit student outcomes. Returning to our example in the beginning of this chapter, one reason for Amanda's relative underperformance compared to Rachel's could be the perceived costs for learning she experiences due to an undiagnosed learning disability. This requires her to put in additional effort to learn the material. Additionally, her fear of failure could be exacerbated by a high-stakes testing environment.

In the above discussion, we have focused on individual types of value and their relationships with educational outcomes. However, in the research literature, it is not uncommon to find general task value scales that are comprised of many different types of task value (e.g., Durik et al., 2006; Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002). This research shows a fairly consistent pattern of relationships: Expectancy beliefs primarily affect performance outcomes, and values generally affect achievement choices (Eccles et al., 1983). For example, Eccles and colleagues (1983) found that math value was a strong predictor of the intention to take more math classes in the future. In a study of science achievement using data from the Trends in International Mathematics and Science Study (TIMSS), science value was related to science interest, school connectedness, and active learning, but unrelated to science achievement (Tighezza, 2013).

10.3.2 Intrinsic and Extrinsic Motivation and Value

Similar to Eccles and colleagues' intrinsic value construct, *intrinsic motivation* is represented in a number of theoretical models as a key reason for valuing an activity. Simply put, intrinsic motivation is defined as the enjoyment of an activity for an activity's sake (Sansone & Harackiewicz, 2000). It reflects engaging in the activity as an end in itself for the inherent pleasure and enjoyment of the activity. Intrinsic motivation is routinely proposed as the optimal reason for an individual to engage in a task. A number of reviews showcase the positive relations between intrinsic motivation and other desirable achievement behaviors and attitudes – in particular that intrinsic motivation is directly related to interest in a task, persisting at the task, and reengaging with the task over time (Lepper & Henderlong, 2000; Sansone & Harackiewicz, 2000). In contrast, *extrinsic motivation* (i.e., valuing a task because it leads to some tangible benefit such as a reward or the avoidance of a punishment) is not well represented in Eccles et al.'s framework.

The contrast between the more controlled reasons for task engagement represented by extrinsic motivation and the more autonomous reasons for task engagement represented by intrinsic motivation is the focus of the motivated regulation

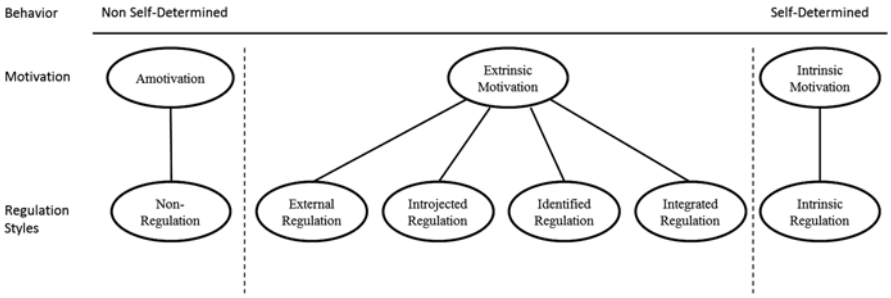

Fig. 10.1 The Motivated Regulation Continuum With Types of Motivation and Regulation Styles (adapted from Deci and Ryan, 2000)

continuum within self-determination theory (Ryan & Deci, 2000). Valuing an activity for extrinsic reasons is on one end of the continuum and valuing an activity for intrinsic reasons is on the other (see Fig. 10.1).

Being extrinsically motivated involves engaging in an activity because of external reward and punishment contingencies and creates compliance to an outside authority. It is considered the least self-determined form of motivation (i.e., the individual need not have any investment in the behavior beyond achieving reward or avoiding punishment). The next form of extrinsic motivation along the self-determination continuum is *introjected motivation*. This category of motivation reflects a shift from responding to external rewards and punishments to internal rewards and punishments. In particular, introjected regulators are still controlled by strong internal pressures that they have certain external obligations that they should or ought to do. When accomplished, ego-related pride is experienced; when unaccomplished, feelings of guilt and anxiety occur. Third on the extrinsic motivation continuum is *identified motivation*. This category reflects an important shift because an individual now sees personal benefits and importance for engaging in the task. This is similar to utility value in the Eccles et al.'s (1983) model. Finally, fourth on the extrinsic motivation continuum is *integrated motivation*. The primary distinction between integrated motivation and true intrinsic motivation is that tasks done to affirm identity (and achieve greater goals) are considered integrated, whereas tasks done for enjoyment are considered intrinsically motivated.

10.3.3 Human Values and Psychological Needs

It is important to define and distinguish values that exist at different conceptual levels. According to Schwartz and Bilsky (1990), *human values* are "beliefs about desirable end states or behaviors that transcend specific situations, guide selection or evaluation of behavior and events, and are ordered by relative importance" (p. 551). These human values can be sorted into two categories, terminal values and instrumental values (Rokeach, 1973). *Terminal values* represent desired end states

that a person or culture holds as important (e.g., social recognition, wisdom, pleasure). *Instrumental values* address modes of conduct to be followed (e.g., honesty, self-control, logic) in order to achieve terminal values. Depending on the theoretical framework, the number of identified human values ranges from 9 (Bilsky, Janik, & Schwartz, 2011) to 36 (Rokeach, 1973).

Self-determination theory also proposes broader human values in the form of three core psychological needs: competence, autonomy, and relatedness. *Competence needs* are fulfilled when an individual has an opportunity to grow, to be efficacious, and to master a task. *Autonomy needs* are met when an individual has choice and feels in control when doing a task. *Relatedness needs* are fulfilled when an individual is able to make meaningful connections to others in a task. When an environment supports an individual's growth on any of these needs, it should hold more value for that individual. This, in turn, is argued to promote that individual's motivation and well-being (Ryan & Deci, 2000). Research on psychological needs reveals that when educators fulfill students' needs for autonomy, relatedness, and competence, they are more intrinsically motivated, regulate their own learning, and perform better (e.g., Niemiec & Ryan, 2009). A number of other psychological needs also have been proposed (Sheldon, Elliot, Kim & Kasser, 2001), such as a need for self-esteem or self-worth (Covington, 1984). This research demonstrates that when achievement is tied to self-worth, thus linking self-esteem to specific levels of performance, then threats to this self-conception (e.g., difficult tasks) undermine achievement, self-regulation, and mental and physical health (e.g., Crocker & Park, 2004).

Both human values and psychological needs are more abstract than task values. *Human* values focus on an individual striving to act a certain way or reach an outcome across situations. In contrast, *task* values focus on the features of a specific task that increase or decrease the relative importance of the task, either compared to other tasks or for attaining an important achievement outcome. Similarly, *psychological needs* appear to be more general – they function as innate values that have the potential to be met in any situation – whereas task values are more specific. To the extent that a task or activity enables a student to meet a psychological need, then the meeting of that need operates as a reason to value the task or activity. To our knowledge, this linkage between needs and values has not been established in the literature, and the field would benefit from further explication of such inter-construct relationships.

10.3.4 Values and Learning Outcomes

Value-related constructs have historically shown moderate to strong associations with achievement choices, task engagement, interest, and achievement (Wigfield & Cambria, 2010). As with the expectancy construct, there have been numerous conceptual reviews of the value construct over the years but no systematic review of the relationship between values and outcomes across student grade level. These reviews

reveal a general decline in mean levels of value across grades (e.g., Jacobs et al., 2002; Lepper, Henderlong, Corpus, & Iyengar, 2005). However, these reviews are silent on how relationships between values and outcomes vary by grade level or change over time.

Instead, two meta-analyses provide empirical support for the relationship between values and student learning outcomes at the college level. In their 2012 meta-analysis of 241 unique data sets, Richardson and colleagues found that valuing education (i.e., academic intrinsic motivation) was positively correlated with college GPA. In their 2004 meta-analysis of 109 studies, Robbins and colleagues found that measures of value (i.e., achievement motivation, academic goals) were predictive of academic performance and persistence in college, even after controlling for socioeconomic status, standardized achievement, and high school GPA.

10.4 Review of Theories that Integrate Expectancy and Value Constructs

By its very name, Eccles' and colleagues' (1983) expectancy-value theory obviously stands out in integrating expectancy and value constructs. Interestingly, Eccles and colleagues shy away from using the term theory. Instead, they refer to their work as an expectancy-value framework or model. As noted at the outset of our chapter, they were motivated to adopt an integrative perspective of various constructs from different motivational theories to better understand students' academic performance, persistence, and choice behaviors. Their framework was also meant to be developmental and contained numerous antecedents of expectancies and values that correspond to terminal and instrumental human values. Elements of the larger expectancy-value framework include the cultural milieu, unique past events, students' perceptions of past events, socializers' behaviors and attitudes, students' perceptions of socializers' attitudes and expectations, and students' goals and self-concept. Thus, in the expectancy-value framework, human values correspond to distal factors in the model of achievement behaviors (e.g., the cultural milieu, student's goals), whereas task values refer to perceptions of the task at hand (i.e., how valuable a task is in attaining a particular goal) (see Fig. 10.2).

Recently, several researchers have noted that the expectancy-value framework promoted by Eccles and colleagues was absent an important element of earlier models. In classic models of achievement motivation, expectancies and values were hypothesized to interact to produce more motivation than either factor alone. In other words, motivation was a product of expectancy times value (i.e., M = E*V). Thus, if either type of motivation was lacking for a given academic task, then a student would not be motivated to engage in it. Using samples from large, international databases, Trautwein and colleagues (2012; Nagengast et al., 2011) found empirical support for the interaction between expectancy and value on engagement in science activities, intentions of pursuing science careers, and academic achievement. These results suggest that students high in both expectancy and value performed better than those high in one or the other or students low in both.

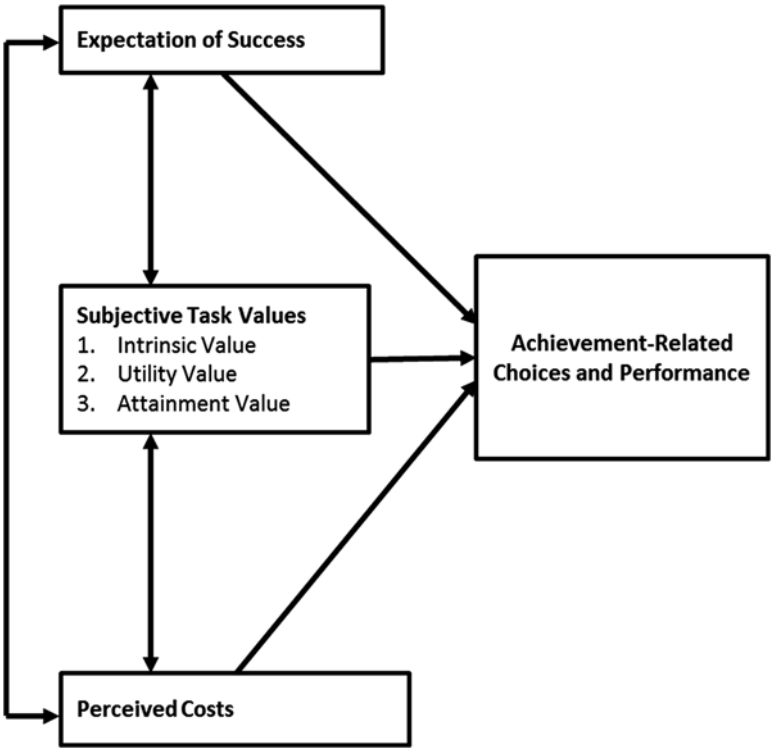

Fig. 10.2 Updated Expectancy-Value-Cost Model of Achievement Motivation (adapted from Eccles et al., 1983; see Barron & Hulleman, 2015)

Besides their own work on expectancy-value models, Eccles and Wigfield (2002) also highlighted numerous other theories that offer integrative perspectives (e.g., Feather, 1988; Weiner, 1992). Although similar in their inclusion of expectancy-related and value-related constructs as sources of student motivation, they vary in focus and motivational mechanisms. For example, self-worth theory includes students' perceptions of their sources of worth and value (Covington, 1984), and control-value theory includes control beliefs (which include both attributions and expectancies) and values as sources of students' achievement emotions (Chap. 11; Pekrun, 2006). More recently, emerging theories of interest also propose an integrative perspective (see Renninger & Hidi, 2011), which defines interest as a combination of expectancy and value (i.e., positive affect, value, and prior knowledge and competence). Our recent theoretical work separating costs from values and expectancies (Barron & Hulleman, 2015) highlights the highly integrative nature of achievement motivation in educational contexts. Instead of being driven by only the positive aspects of task engagement (e.g., success expectancies and task values), students are often mindful of the obstacles to engagement and potential negative affect they will experience.

10.5 Intervention Drivers: Research-Based Sources of Students' Expectancies, Values, and Costs

As our review of the literature demonstrates, an expectancy-value framework serves as a conceptual umbrella under which other motivation theories and constructs can easily fit. Additionally, we have proposed a revised framework that includes cost as a separate, third component (Barron & Hulleman, 2015). To be optimally motivated, students need to have expectancy beliefs that they can be successful in their schoolwork (i.e., *Do students think they can do the task?*) and see value for their schoolwork (i.e., *Do students want to do the task?*). However, even if students believe they can do a task and have a reason to do a task, they still might not be motivated if they experience significant cost preventing them from engaging in that task (i.e., *Are students free of barriers preventing them from investing time, energy, or resources into the task?*). Thus, as teachers encounter motivation problems with their students, deciding if the problem is an expectancy, a value, or a cost problem is a critical first step in determining how to intervene.

Therefore, based on our review of expectancy, value, and cost constructs within educational and social psychology, and a desire to identify pathways for practitioners to enhance student motivation, we have identified research-based sources of expectancy, value, and cost that are potentially amenable to interventions (see Tables 10.3, 10.4, and 10.5, respectively). These sources refer to the underlying psychological processes that both serve as antecedents of expectancy, value, or cost and that are potentially amenable to intervention by educational practitioners, including teachers, parents, and administrators. Importantly, these sources can serve as the targets or drivers of interventions aimed at enhancing student outcomes by boosting students' expectancies and values and reducing their costs. Although there are additional sources of expectancies, values, and costs – such as those identified in the Eccles model, including cultural milieu and socializers' goals and expectations – the sources in our tables have been identified as being the most accessible to change through direct intervention.

We therefore conclude this chapter with a brief introduction to a growing body of intervention work designed to promote student learning outcomes by targeting sources of students' expectancy, value, and costs for their schoolwork.

10.6 Interventions and Programs that Foster Motivation

Much of the work that we have reviewed above was based on correlational research that links self-report measures of expectancy, value, and cost to student outcomes. Although observational and correlational research can generate and test hypotheses, intervention research (i.e., research that formally manipulates an independent variable) provides valuable information about what happens when we attempt to enhance educational outcomes through intentional manipulation. From a theoretical

Table 10.3 Research-based sources of expectancy-related beliefs

Expectancy source	Definition
Perceptions of ability/skill	When students perceive they have a high level of ability and/or skill at an activity, they are more likely to experience high expectancy (Bandura, 1997; Wigfield & Eccles, 2002).
Effort attributions	When students believe that their effort will lead to learning, they are more likely to experience high expectancy (Dweck & Leggett, 1988; Dweck, 1999; Weiner, 1972).
Success experiences	When students are successful at an activity, or watch others have success, they are more likely to experience high expectancy (Bandura, 1997; Eccles et al., 1983).
Support and scaffolding	When students are appropriately supported in completing an activity (e.g., through encouragement and having the resources necessary to complete the task), they are more likely to experience high expectancy (Bandura, 1997).
Clear expectations	When students know what is expected of them on an activity, and have clearly defined goals, they are more likely to experience high expectancy (Pajares, 1996).
Appropriate challenge	When the difficulty of the task or activity matches students' skill levels, they are more likely to experience high expectancy (Eccles et al., 1983).
Feedback	When students receive feedback that effort matters and skills are amenable to change and are task focused (rather than ability focused), they are more likely to experience high expectancy (Dweck & Leggett, 1988; Dweck, 1999).
Growth experiences	When students engage in learning activities that challenge them to grow and learn, and experience growth in their skills and performance improvements, they are more likely to experience both high expectancy and value (Dweck & Leggett, 1988; Dweck, 1999; Hong et al., 1999).
Perceptions of others' expectations	Parents' and teachers' expectancies and attitudes shape children'/students' expectancies; for instance, if teachers have high expectations for their students, these students in turn develop high expectancies (Bandura, 1997; Dweck & Leggett, 1988; Dweck, 1999; Eccles et al., 1983).
Perceived task difficulty	When students perceive a subject or task as being not difficult, they develop higher estimates of their own abilities for the subject or task (Bandura, 1997; Pajares, 1996; Wigfield & Eccles, 2002).
Stability attributions	When students attribute success to a stable factor (ability), then they will have higher expectations for future success; if they attribute it to an unstable factor (good luck), they will be uncertain about future success and have lower expectations for future success (Weiner, 2010).

perspective, intervention research helps move the field forward by providing insight about the causal relationships between motivation constructs and educational outcomes or between educational settings and motivation outcomes (Shadish, Cook, & Campbell, 2002; Tunnell, 1977). Because interventions represent the operationalized theory in action, they provide a strong test of the theory as applied in an educational context.

Table 10.4 Research-based sources of value

Value source	Definition
Intrinsic benefits	When students find the activities and academic content enjoyable and interesting, they are more likely to experience high value (Renninger & Hidi, 2011).
Relevance	When students are able to connect what they are learning to their personal lives and/or the real world, they are more likely to experience high value (Hulleman & Harackiewicz, 2009).
Context and rationale	When students understand that an activity is meaningful and has a purpose, they are more likely to experience high value (Lepper & Henderlong, 2000).
Variety and novelty	When students engage in activities that are varied and novel, they are more likely to experience high value (e.g., catch and hold interest; Hidi & Renninger, 2006).
Enthusiastic models	When students interact with teachers and other adults who are enthusiastic and passionate about learning, they are more likely to experience high value (Patrick, Hisley, & Kempler, 2000).
Growth experiences	When students engage in learning activities that challenge them to grow and learn, and experience growth in their skills and performance improvements, they are more likely to experience both high expectancy and value (Dweck & Leggett, 1988; Dweck, 1999; Hong et al., 1999)
Choice and control	When students feel a sense of control and choice over their learning, they are more likely to experience high value (Patall et al., 2010).
Positive relationships and sense of belongingness	When students experience meaningful student-student and student-teacher relationships, they are more likely to experience high value (Furrer & Skinner, 2003; Walton & Cohen, 2007).
Extrinsic benefits	When students receive external rewards and incentives for learning (e.g., prizes, food), they are more likely to experience high value to complete an activity but low value to produce quality work (Marinak & Gambrell, 2008).

Table 10.5 Research-based sources of cost

Value source	Definition
Effort and time needed for the activity	When students feel that the workload is unreasonable (e.g., 5 hours/night) and/or unnecessary (e.g., busy work), they are more likely to experience increased cost (Parsons et al., 1980; Perez et al., 2014).
Effort and time needed for other competing activities	When student have too many other demands on their time or do not know how to effectively manage their time, they are more likely to experience high cost (Barron & Hulleman, 2015; Flake et al., 2015).
Loss of valued alternatives	When students feel like the learning activity is not worth their time compared to other things they might do (e.g., socializing), they are more likely to experience high cost (Conley, 2012; Perez et al., 2014).
Psychological and physical reactions to the activity	When students feel unsafe and uncomfortable, either physically or psychologically (e.g., nervous, bored, tired), they are more likely to experience high cost (Eccles et al., 1983; Ramirez & Beilock, 2011).

From a practical perspective, intervention studies facilitate our understanding about which interventions are most effective in improving educational outcomes in a way that observational research cannot. This understanding can guide recommendations for educational practice based on appropriate scientific evidence. For example, what is the best way to prevent students like Amanda from disengaging in the learning process? It is not enough simply to know that some motivation constructs are correlated with important student outcomes. What is needed are interventions designed to target motivational constructs and processes that, in turn, enhance educational outcomes.

Although several narrative reviews have highlighted important constructs (Pintrich, 2003) and interventions (Wigfield & Wentzel, 2007; Yeager & Walton, 2011) that are linked to enhanced student motivation and outcomes, we were interested in finding interventions that had the strongest empirical support as implemented within actual educational contexts, as opposed to correlational or laboratory studies. To that end, we recently conducted a meta-analysis of motivation interventions conducted in ecologically valid[1] educational contexts (Lazowski & Hulleman, 2013). As presented in Table 10.6, we found over 63 different interventions designed to enhance student motivation in education contexts. Although these interventions originate from 12 different theoretical frameworks, we were able to categorize these interventions as targeting expectancy-related, value-related, or cost-related sources. We found that these 63 interventions produced an average effect size on behavioral, self-report, and performance outcomes of two-thirds of a standard deviation ($d = 0.58$). Below, we highlight some examples that have the strongest empirical support.

10.6.1 Expectancy Interventions

Attribution Retraining One set of interventions aimed at changing students' success expectancies has focused primarily on changing cognitive attributions for success and failure. Many of these interventions provide students with training about ascribing academic success to things that are within their control (e.g., effort) and that academic difficulties can be overcome. These control-enhancing interventions have been successful in increasing perceived academic control, and these changes mediate effects on academic motivation and achievement outcomes (e.g., Hall,

[1] We used Tunnell's (1977) three dimensions of naturalness to help define ecologically valid: *natural treatments* are naturally occurring events to which the participant is exposed (e.g., pedagogical practices, curriculum); *natural settings* are those that are not perceived to be established for the purposes of research (e.g., a non-laboratory setting); and *natural behavior* occurs on its own within the educational context (e.g., statewide mandated standardized tests will be taken by students whether they are in a study or not). Intervention studies that contain these dimensions of naturalness are more likely to have results that will generalize to other settings; therefore, we selected studies that contained at least one dimension of naturalness and that targeted student motivation.

Table 10.6 Summary table of motivation intervention studies by research-based source of expectancy, value, and cost

Study	Source	Avg. d^c	n_e, n_c[a]	Age[b]
Expectancy interventions				
Boese et al. (2013)	Ability/skill Effort	0.77	84, 42	C
Hall et al. (2007)	Ability/skill Effort	0.31	374, 375	C
Hall et al. (2004)	Ability/skill Etffort	0.35	101, 102	C
Ruthig et al. (2004)	Ability/skill Effort	0.61	118, 118	C
Struthers and Perry (1996)	Ability/skill Effort	0.41	108, 150	C
Wilson and Linville (1985)	Ability/skill Effort	0.36	20, 20	C
Wilson and Linville (1982)	Ability/skill Effort	0.73	20, 20	C
Yeager et al. (2013) Study 1 Study 2 Study 3	Ability/skill Effort	0.76 0.78 0.44	22, 22 22, 22 38, 38	MS MS HS
Craven, Marsh, and Debus (1991)	Ability/skill Effort/feedback	0.08	81, 79	ES
Aronson et al. (2002)	Effort	0.57	37, 37	C
Blackwell et al. (2007)	Effort	0.69	49, 50	MS
Good et al. (2003)	Effort	0.92	69, 69	MS
Mueller and Dweck (1998)	Effort			
Study 1		0.84	64, 64	ES
Study 2		1.17	25, 26	ES
Study 3		0.81	44, 44	ES
Study 4		1.15	25, 26	ES
Study 5		1.03	23, 23	ES
Study 6		1.28	24, 24	ES
Gollwitzer and Brandstatter (1997)	Challenge Feedback	1.24	43, 43	C
Kitsantas et al. (2004)	Challenge Feedback	0.73	48, 48	HS
Morisano et al. (2010)	Challenge Feedback	0.44	43, 42	C
Muis, Ranellucci, Franco, and Crippen (2013)	Challenge Feedback	0.12	198, 52	C
Silva, White, and Yoshida (2011)	Challenge	0.71	20, 21	HS
Hofer and Yu (2003)	Support and scaffolding	0.48	39, 39	C

(continued)

Table 10.6 (continued)

Study	Source	Avg. d^c	n_e, n_c[a]	Age[b]
Greenstein (1976)	Feedback	0.54	87, 84	C
Duckworth, Kirby, Gollwitzer, and Oettingen (2015)	Mental contrasting Challenge	0.51	38, 39	ES
Value interventions				
Acee and Weinstein (2010)	Relevance Context	0.56	41, 41	C
Harackiewicz et al. (2012)	Relevance	0.32	94, 94	HS
Hulleman et al. (2010), Study 2	Relevance	0.38	160, 158	C
Hulleman and Harackiewicz (2009)	Relevance	0.28	136, 126	HS
Patall et al. (2010)	Intrinsic Choice/control	0.12	193, 194	HS
Vansteenkiste, Timmermans, Lens, Soenens, and Van den Broeck (2008)	Intrinsic Choice/control	0.70	68, 70	MS
Vansteenkiste et al. (2005)	Intrinsic Choice/control			
Study 1		0.83	65, 65	MS
Study 2		0.74	57, 56	MS
Vansteenkiste et al. (2004) Study 1 Study 2	Intrinsic Choice/control	1.57 1.49	100, 100 189, 189	C C
Vansteenkiste et al. (2004)	Context Growth	0.47	123, 122	C
Hoyert and O'Dell (2006)	Context Growth	1.08	69, 68	C
Miller and Meece (1997)	Context Growth	0.54	94, 93	ES
Guthrie et al. (2006)	Variety/novelty	0.71	49, 49	ES
Hidi et al. (2002)	Variety/novelty	0.67	90, 90	MS
Schaffner and Schiefele (2007)	Intrinsic	0.46	188, 187	HS
Day et al. (1994)	Context	0.91	42, 41	ES
Oyserman et al. (2002)	Context	0.32	62, 146	MS
Cook, Purdie-Vaughns, Garcia, and Cohen (2012)	Self-affirmation	0.36	61, 60	MS
Miyake et al. (2010)	Self-affirmation	0.21	69, 47	C
Sherman et al. (2013) Study 1 Study 2	Self-affirmation	0.34 0.64	41, 40 26, 29	MS MS
Cost interventions				
Cohen et al. (2006)	Psychological	0.75	121, 122	MS
Cohen et al. (2009)	Psychological	0.51	192, 193	MS
Sherman et al. (2013), Study 1	Psychological	0.36	41, 40	MS

(continued)

Table 10.6 (continued)

Study	Source	Avg. d^c	n_e, n_c[a]	Age[b]
Expectancy and value interventions				
Guthrie et al. (2000)	Growth Scaffolding Belongingness Relevance	0.67	79, 83	ES
Garcia and De Caso (2004)	Effort Relevance Scaffolding	0.47	66, 61	ES
Martin (2008)	Effort Scaffolding	0.48	26, 27	HS
Froiland (2011)	Growth Scaffolding Choice/control	0.73	15, 15	ES
Reeve et al. (2004)	Growth Scaffolding Choice/control	1.69	10, 10	HS
Value and cost interventions				
Hausmann, Ye, Schofield, and Woods (2009)	Belongingness Psychological	0.26	70, 67	C
Walton and Cohen (2007)	Belongingness Psychological	1.03	81, 81	C
Walton and Cohen (2011)	Belongingness Psychological	0.58	49, 43	C
Expectancy and cost interventions				
Jamieson, Mendes, Blackstock, and Schmader (2010)	Attribution Challenge	0.87	14, 14	C
Total		**0.66**[d]	**4738, 4634**	

[a]The sample size for the experimental condition (n_e) is reported first, followed by the sample size for the control condition (n_c).
[b]Grade levels included elementary school (ES), middle school (MS), high school (HS), and college (C).
[c]Types of dependent variables included self-report (SR), behavioral indicator (B), and performance indicator (P).
[d]For more details, see Lazowski and Hulleman (2015).

Hladkyj, Perry, & Ruthig, 2004; Haynes, Ruthig, Perry, Stupnisky, & Hall, 2006; Perry, Stupnisky, Hall, Chipperfield, & Weiner, 2010).

There have been several studies demonstrating that changes in causal attributions relate to changes in academic achievement. Many of these intervention studies sought to alter the attributions that low performing students made regarding their academic achievement from low ability to underscoring the importance of effort and the notion that achievement was amenable to change. These shifts in attribution have been demonstrated to improve course grades (Boese, Stewart, Perry, & Hamm,

2013; Hall et al., 2007, 2004; Yeager, Paunesku, Walton, & Dweck, 2013), exam performance (Struthers & Perry, 1996), GPA (Boese et al., 2013; Ruthig, Perry, Hall, & Hladkyj, 2004; Yeager et al., 2013; Wilson & Linville, 1982, 1985), standardized test scores (Good et al., 2003; Wilson & Linville, 1982, 1985), intrinsic motivation (Hall et al., 2007), and reduction in text anxiety and voluntary course withdrawal (Ruthig et al., 2004).

Growth Mindsets Based on Dweck's theory of the malleability of intelligence, the growth mindset intervention targets students' perceptions about their capacity to learn. There have been several versions of the growth mindset intervention that have been demonstrated to be effective in enhancing student outcomes. Blackwell and colleagues (2007) developed eight, 1-hour sessions for middle school students. Six of the sessions instructed students on the latest research on how the brain develops and grows. Two additional sessions focused on helping students understand that their brains can grow through effort and persistence through difficulty and using appropriate learning strategies. Students who were randomly assigned to the mindset intervention had higher self-reported motivation and academic performance compared to those in the control condition. Other versions of the intervention have replicated this effect in high school and college students (Aronson et al., 2002; Yeager et al., 2013).

10.6.2 Value Interventions

Utility Value Based on Eccles' expectancy-value framework, Hulleman and colleagues developed and tested interventions designed to increase students' perceptions of the relevance of academics to their lives (i.e., utility value). In one set of studies, students were randomly assigned to either write about how the course material related to their lives or write a summary of the material they were studying. The findings revealed that high school science students (Hulleman & Harackiewicz, 2009), college psychology students (Hulleman et al., 2010), college biology students (Harackiewicz, Canning, Tibbetts, Priniski, & Hyde, 2015), and statistics students (Hulleman, An, Hendricks, & Harackiewicz, 2007) in the utility value treatment condition reported greater topic interest, future intentions, and academic performance than students in the control condition. These effects were particularly strong for students with low actual or expected academic performance. In another study, parents of high school students were randomly assigned to receive information that highlighted the utility value of mathematics and science courses for their teenagers, along with strategies on how to talk to their teenagers about the value of math and science coursework. Students whose parents received the information took more mathematics and science courses in their last 2 years of high school than students whose parents did not receive the information (Harackiewicz, Rozek, Hulleman, & Hyde, 2012).

Choice Several interventions have demonstrated the impact of increasing value through opportunities for choice in the classroom. In one study, Patall, Cooper, and Wynn (2010) randomly assigned high school students to receive a choice of homework assignments or no choice. Students in the choice condition had higher self-reported intrinsic motivation and perceived competence, and also performed better on the unit exam, than students in the no-choice condition. Vansteenkiste, Simons, Lens, Sheldon, and Deci (2004) randomly assigned college students to conditions that appeared to have more or less choice. The perceived-choice condition boosted students' depth of processing, persistence, and test performance compared to the no-choice condition.

10.6.3 Cost Interventions

Values Affirmation The emotional cost of academic life can manifest itself when students identify with groups of students who are stereotyped to under-perform. This perceived threat, known as stereotype threat (Steele, 1997), can undermine academic performance and persistence, resulting in a sorting mechanism that reduces minority success and completion rates in high school and college. An intervention designed to ameliorate this perceived threat has been developed and tested by Geoffrey Cohen and colleagues (e.g., Cohen, Garcia, Apfel, & Master, 2006; Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009). Students randomly assigned to the affirmation condition wrote about their top most important values, whereas students assigned to the control condition wrote about their least important values. By writing about their most important values, students are affirming core aspects of themselves, and this affirmation serves as a buffer against threats in a single domain. The results of this brief intervention are startling. In a sample of seventh grade students, the values-affirmation intervention reduced the black-white achievement gap by 40 % (Cohen et al., 2006). In a 2-year follow-up, the benefits of the intervention were particularly acute for low-achieving black students who increased their performance relative to the control group (Cohen et al., 2009). This intervention effect has been replicated with other minority groups, such as Latino American middle school students (Hanselman, Bruch, Gamoran, & Borman, 2014; Sherman et al., 2013) and first-generation college students (Harackiewicz et al., 2014).

Belonging In addition to stereotype threat, students can also experience emotional cost in an academic setting if they feel anxious about not belonging or fitting in with other students. These feelings of belonging uncertainty can lead to students withdrawing from the academic experience and result in poorer learning and health outcomes (Walton & Cohen, 2007). In a series of studies, Greg Walton and Geoffrey Cohen developed an intervention targeting students' feelings of belonging in the academic environment (Walton & Cohen, 2007, 2011). Students randomly assigned to the intervention condition read results of a survey and quotes from other students

that emphasized that everyone struggles with some aspects of college initially and that these initial difficulties were temporary. In essence, students learned that there were other students like them who initially felt like they did not fit in but who eventually succeeded in college. The results indicated students most likely to feel uncertain about belonging in college – African American students – demonstrated increased GPA and self-reported health and well-being (2011).

10.6.4 Multicomponent Interventions

Thus far, we have reviewed interventions that target a single motivational construct or component. However, it is possible that to be maximally effective interventions need to address multiple facets of the student experience. These interventions could target multiple motivational constructs, or these interventions could include pedagogical elements that target particular types of learning, such as reading or mathematics. As a group, such multicomponent interventions have received less experimental evaluation in the literature, so the associated empirical base is not as strong. Below, we review two promising multicomponent interventions in the literature that specifically target motivational processes to enhance student learning outcomes.

An intervention developed by Andrew Martin provides an example of a multicomponent motivation intervention. Designed using an integrative motivation and engagement framework known as The Wheel (Martin, 2008), this intervention targets students' adaptive and maladaptive behaviors and cognitions. Delivered over the course of 13 modules, students are guided through instruction on the 11 aspects of the wheel: self-efficacy and mastery (expectancy); valuing (value); anxiety, failure avoidance, uncertain control, self-handicapping, and disengagement (cost); and persistence, planning, and task management (learning skills). Initial quasi-experimental results indicate that the intervention boosted students' self-reported motivation and persistence (Martin, 2008).

The Concept-Oriented Reading Instruction (CORI) intervention is an example of a multicomponent intervention that combines motivational aspects with reading strategy instruction. Developed by John Guthrie and Allan Wigfield, CORI links reading fiction and nonfiction books to science activities (Guthrie, Wigfield, & VonSecker, 2000). This reading program is organized into thematic units designed to target five motivational processes: self-efficacy and mastery goals (expectancy), perceived autonomy and intrinsically motivating activities (value), and collaborative work that provides social support for learning (cost) (Guthrie, McRae, & Lutz Klauda, 2007). A meta-analysis of 11 quasi-experimental studies demonstrates that the CORI intervention improves students' reading strategy use, self-reported reading motivation, and achievement (Guthrie et al., 2007).

Certainly, examples of additional multicomponent interventions abound in the literature. In a special issue of the *Educational Psychologist* edited by Allan Wigfield and Kathryn Wentzel, the authors of different articles discuss school-wide

reform efforts to create positive social and emotional climates for children (Juvonen, 2007), small learning community reforms (Felner, Seitsinger, Brand, Burns, & Bolton, 2007), and social skills training for aggressive children (Hudley, Graham, & Taylor, 2007). Such interventions connect to broader literature on social-emotional interventions (e.g., Collaborative for Academic, Social, and Emotional Learning, 2013; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011) and may indirectly target motivation through instruction in social and emotional skills (Rimm-Kaufman & Hulleman, in press), such as emotional regulation and decision-making (see Chap. 13).

10.7 Caution: One Size Does Not Fit All

It is important to note that these interventions are not "magic bullets" that can work for all students in all situations (cf. Durik, Hulleman, & Harackiewicz, 2014; Yeager & Walton, 2011). If students already have growth mindsets, then a growth mindset intervention may not be of benefit. However, if students are worried that they may not be able to find friends as they transition from high school to college, then a belonging intervention may buffer their concerns and facilitate their engagement in their academics. In addition, psychological interventions target specific mechanisms that, if not implemented properly or if used inappropriately, can have unintended negative consequences. For example, a common reaction to an apparent lack of student motivation is to offer rewards to students (see Table 10.4). But without knowledge of the reasons for a lack of engagement, provision of rewards may not produce the desired result (Marinak & Gambrell, 2008).

On the one hand, being offered financial compensation for each "A" earned can provide students a reason to value learning, particularly when the student lacks any other value for the activity. In this case, when students see no reason to engage in an activity, then rewards might instigate some engagement in the activity. On the other hand, being motivated by extrinsic reasons can lead to suboptimal outcomes. A host of correlational evidence demonstrates that students' self-reports of extrinsic, compared with intrinsic, motivation is negatively related to outcomes (e.g., Lepper et al., 2005; Vallerand et al., 1993). For example, students who report having higher quality of motivation (high intrinsic, low extrinsic) have higher achievement than students with higher quantity of motivation (high intrinsic, high extrinsic; Hayenga & Henderlong Corpus, 2010). Furthermore, experimental evidence indicates that tangible, extrinsic rewards can undermine students' motivation to engage in academic tasks, particularly if the rewards are unrelated to future task engagement (e.g., Marinak & Gambrell, 2008) and are perceived as controlling or are expected (e.g., Deci, Koestner, & Ryan, 1999, 2001; Tang & Hall, 1995; but see Cameron, 2001; Cameron & Pierce, 1994). The conclusion, besides the fact that rewards are complicated, is that knowledge of the underlying motivational issue is vitally important before implementing any of these interventions as potential solutions to a lack of student engagement.

10.8 Conclusion

From a broader perspective, the conceptual models of expectancy- and value-related constructs, whether integrative or singularly focused, attempt to identify the antecedents and sources of expectancies and values, delineate how expectancies and values develop over time, and determine their contribution to student learning outcomes and success. As we have reviewed, students' expectancy and value beliefs are central predictors of educational outcomes and attainment. In addition, having the skills to learn and persist in the face of challenging academic tasks is central to students' future success, whether it be in attaining educational credentials, choosing a career path, or maintaining long-term employment. Thus, if we are to leverage the relationships between expectancy-value motivation and learning outcomes, it is critical to identify the sources of expectancy and value that are malleable and potentially accessible to educational practitioners. By targeting motivation gaps, educational practitioners, policy-makers, and researchers have a potentially powerful tool to further close achievement gaps and inspire more students to persist academically, both in the short and long term.

References

Acee, T. W., & Weinstein, C. E. (2010). Effects of a value-reappraisal intervention on statistics students' motivation and performance. *The Journal of Experimental Education, 78*, 487–512.

Ames, R. (1983). Help-seeking and achievement orientation, perspectives from attribution theory. In B. DePaulo, A. Nadler, & J. Fisher (Eds.), *New directions in helping: Vol. 2. Help-seeking* (pp. 165–186). New York: Academic.

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African-American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38*(2), 113–125.

Atkinson, J. W. (1958). Towards experimental analysis of human motivation in terms of motives, expectancies, and incentives. In J. W. Atkinson (Ed.), *Motives in fantasy, action, and society* (pp. 288–305). Princeton, NJ: D. Van Nostrand Company.

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist, 44*(9), 1175–1184.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117–148.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Baranik, L. E., Barron, K. E., & Finney, S. J. (2010). Specificity of achievement goal measures: An empirical investigation comparing general vs. specific measures. *Human Performance, 23*, 155–172.

Barron, K. E., & Hulleman, C. S. (2015). The expectancy-value-cost model of motivation. In J. D. Wright (Ed.), *International encyclopedia of the social and behavioral sciences, 2nd edition* (Vol. 8, pp. 503–509). Oxford: Elsevier Ltd.

Battle, A., & Wigfield, A. (2003). College women's value orientations toward family, career, and graduate school. *Journal of Vocational Behavior, 62*(1), 56–75.

Bilsky, W., Janik, M., & Schwartz, S. H. (2011). The structural organization of human values – Evidence from three rounds of the European Social Survey (ESS). *Journal of Cross-Cultural Psychology, 42*, 759–776.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*(1), 246–263.

Bodill, K., & Roberts, L. D. (2013). Implicit theories of intelligence and academic locus of control as predictors of studying behaviour. *Learning and Individual Differences, 27*, 163–166.

Boese, G. D., Stewart, T. L., Perry, R. P., & Hamm, J. M. (2013). Assisting failure-prone individuals to navigate achievement transitions using a cognitive motivation treatment (attributional retraining). *Journal of Applied Social Psychology, 43*(9), 1946–1955.

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*(1), 1–40.

Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin, 139*(3), 655–701.

Byrne, B. M., & Shavelson, R. J. (1986). On the structure of adolescent self-concept. *Journal of Educational Psychology, 78*(6), 474–481.

Cameron, J. (2001). Negative effects of reward on intrinsic motivation – A limited phenomenon: Comment on Deci, Koestner, and Ryan (2001). *Review of Educational Research, 7*(1), 29–42.

Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research, 64*(3), 363–423.

Ciarrochi, J., Heaven, P. C., & Davies, F. (2007). The impact of hope, self-esteem, and attributional style on adolescents' school grades and emotional well-being: A longitudinal study. *Journal of Research in Personality, 41*(6), 1161–1178.

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science, 313*(5791), 1307–1310.

Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science, 324*(5925), 400–403.

Collaborative for Academic, Social, and Emotional Learning. (2013). *Effective social and emotional learning programs: Preschool and elementary school edition*. Collaborative for Academic, Social and Emotional Learning. Retrieved from http://www.casel.org/guide

Conley, A. M. (2012). Patterns of motivation beliefs: Combining achievement goal and expectancy-value perspectives. *Journal of Educational Psychology, 104*(1), 32–47.

Connell, J. P., & Wellborn, J. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In M. Gunnar & A. Sroufe (Eds.), *Minnesota symposium on child development* (Vol. 22, pp. 43–77). Hillsdale, MI: Erlbaum.

Cook, J. E., Purdie-Vaughns, V., Garcia, J., & Cohen, G. L. (2012). Chronic threat and contingent belonging: Protective benefits of values affirmation on identity development. *Journal of Personality and Social Psychology, 102*(3), 479–496.

Covington, M. V. (1984). The self-worth theory of achievement motivation: Findings and implications. *Elementary School Journal, 85*, 5–20.

Crandall, V. J., Katkovsky, W., & Preston, A. (1962). Motivational and ability determinants of young children's intellectual achievement behaviors. *Child Development, 33*, 643–661.

Craven, R. G., Marsh, H. W., & Debus, R. L. (1991). Effects of internally focused feedback and attributional feedback on enhancement of academic self-concept. *Journal of Educational Psychology, 83*(1), 17–27.

Crocker, J., Luhtanen, R. K., Cooper, M. L., & Bouvrette, A. (2003). Contingencies of self-worth in college students: theory and measurement. *Journal of Personality and Social Psychology, 85(5)*, 894–908.

Crocker, J., & Park, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin, 130*(3), 392–414.

Day, J. D., Borkowski, J. G., Punzo, D., & Howsepian, B. (1994). Enhancing possible selves in Mexican American students. *Motivation and Emotion, 18*, 79–103.

De Volder, M. L., & Lens, W. (1982). Academic achievement and future time perspective as a cognitive–motivational concept. *Journal of Personality and Social Psychology, 42*(3), 566–571.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*(6), 627–668.

Deci, E. L., Koestner, R., & Ryan, R. M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research, 71*(1), 1–27.

Denissen, J. J., Zarrett, N. R., & Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development, 78*(2), 430–447.

Duckworth, A. L., Kirby, T., Gollwitzer, A., & Oettingen, G. (In press). From fantasy to action: Mental contrasting with implementation intentions (MCII) improves academic performance in children. *Social Psychological and Personality Science, 4*(6), 745–753.

Durik, A. M., Hulleman, C. S., & Harackiewicz, J. M. (2014). One size fits some: Instructional enhancements to promote interest don't work the same for everyone. In K. A. Renninger & M. Nieswandt (Eds.), *Interest, the self, and K-16 mathematics and science learning*. Washington, DC: American Educational Research Association.

Durik, A. M., Vida, M., & Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: A developmental analysis. *Journal of Educational Psychology, 98*(2), 382–393.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*(1), 405–432.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Taylor and Francis/Psychology Press.

Dweck, C. (2006). *Mindset: The new psychology of success*. New York: Random House LLC.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256–273.

Dweck, C. S., Walton, G. M., & Cohen, G. L. (2011). *Academic tenacity* (White paper prepared for the Gates Foundation). Seattle, WA: Gates Foundation.

Eccles, J. (1980). *Self-perceptions, task-perceptions and academic choice: Origins and change* (Final Report). Ann Arbor, MI: The University of Michigan, Department of Psychology.

Eccles, J. (2005). Subjective task value and the Eccles et al. model of achievement-related choices. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 105–121). New York: Guildford.

Eccles, J. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist, 44*(2), 78–89.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 74–146). San Francisco: W.H. Freeman.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*, 109–132.

Eccles, J. S., Wigfield, A., Harold, R., & Blumenfeld, P. B. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development, 64*, 830–847.

Feather, N. T. (1988). Values, valences, and course enrollment: Testing the role of personal values within an expectancy-valence framework. *Journal of Educational Psychology, 80*(3), 381–391.

Felner, R. D., Seitsinger, A., Brand, S., Burns, A., & Bolton, N. (2007). Creating small learning communities: Lessons from the project on high-performing learning communities about "what works" in creating productive, developmentally enhancing, learning contexts. *Educational Psychologist, 42*, 209–221.

Finney, S. J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology, 28*(2), 161–186.

Flake, J., Barron, K. E., Hulleman, C. S., McCoach, D. B., & Welsh, M. E. (2015). Measuring cost: The Forgotten component of expectancy-value theory. *Contemporary Educational Psychology, 41*, 232–244.

Froiland, J. M. (2011, April). Parental autonomy support and student learning goals: A preliminary examination of an intrinsic motivation intervention. In C. F. Weems (Ed.), *Child & Youth Care Forum* (Vol. 40, No. 2, pp. 135–149). New York, NY: Springer.

Fulmer, S. M., & Tulis, M. (2013). Changes in interest and affect during a difficult reading task: Relationships with perceived difficulty and reading fluency. *Learning and Instruction, 27*, 11–20.

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology, 95*(1), 148–162.

Garcia, J. N., & de Caso, A. M. (2004). Effects of a motivational intervention for improving the writing of children with learning disabilities. *Learning Disability Quarterly, 27*(3), 141–159.

Gaspard, H., Dicke, A., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through higher differentiation: Gender differences in value beliefs about math, *Journal of Educational Psychology, 107*(3), 663–677.

Glasgow, K. L., Dornbusch, S. M., Troyer, L., Steinberg, L., & Ritter, P. L. (1997). Parenting styles, adolescents' attributions, and educational outcomes in nine heterogeneous high schools. *Child Development, 68*(3), 507–529.

Goetz, T., & Bieg, M. (2016). Academic emotions and their regulation via emotional intelligence. In A. A. Lipnevich, F. Preckel, & R. D. Roberts (Eds.), *Psychosocial skills and school systems in the twenty-first century: Theory, research, and applications (pp. xx–xx)*. Cham: Springer.

Gollwitzer, P. M., & Brandstätter, V. (1997). Implementation intentions and effective goal pursuit. *Journal of Personality and Social Psychology, 73*(1), 186–199.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology, 24*(6), 645–662.

Grays, M. P. (2013). *Measuring motivation for coursework across the academic career: A longitudinal invariance study*. Unpublished doctoral dissertation, James Madison University, Harrisonug, VA.

Greenstein, T. (1976). Behavior change through value self-confrontation: A field experiment. *Journal of Personality and Social Psychology, 34*, 254–262.

Guthrie, J. T., McRae, A., & Klauda, S. L. (2007). Contributions of concept-oriented reading instruction to knowledge about interventions for motivations in reading. *Educational Psychologist, 42*(4), 237–250.

Guthrie, J. T., Wigfield, A., Humenick, N. M., Perencevich, K. C., Taboada, A., & Barbosa, P. (2006). Influences of stimulating tasks on reading motivation and comprehension. *The Journal of Educational Research, 99*, 232–246.

Guthrie, J. T., Wigfield, A., & VonSecker, C. (2000). Effects of integrated instruction on motivation and strategy use in reading. *Journal of Educational Psychology, 92*(2), 331.

Hall, N. C., Hladkyj, S., Perry, R. P., & Ruthig, J. C. (2004). The role of attributional retraining and elaborative learning in college students' academic development. *The Journal of Social Psychology, 144*(6), 591–612.

Hall, N. C., Perry, R. P., Goetz, T., Ruthig, J. C., Stupnisky, R. H., & Newall, N. E. (2007). Attributional retraining and elaborative learning: Improving academic development through writing-based interventions. *Learning and Individual Differences, 17*(3), 280–290.

Haney, P., & Durlak, J. A. (1998). Changing self-esteem in children and adolescents: A meta-analytical review. *Journal of Clinical Child Psychology, 27*(4), 423–433.

Hanselman, P., Bruch, S. K., Gamoran, A., & Borman, G. D. (2014). Threat in context: School moderation of the impact of social identity threat on racial/ethnic achievement gaps. *Sociology of Education, 87*(2), 106–124.

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Giffen, C. J., Blair, S. S., Rouse, D. I., et al. (2014). Closing the social class achievement gap for first-generation students in undergraduate biology. *Journal of Educational Psychology, 106*(2), 375–389.

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2015). Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *Journal of Personality and Social Psychology, 2*. Online first edition.

Harackiewicz, J. M., Rozek, C. S., Hulleman, C. S., & Hyde, J. S. (2012). Helping parents to motivate adolescents in mathematics and science: An experimental test of a utility-value intervention. *Psychological Science, 23*(8), 899–906.

Hausmann, L. R., Ye, F., Schofield, J. W., & Woods, R. L. (2009). Sense of belonging and persistence in White and African American first-year students. *Research in Higher Education, 50*(7), 649–669.

Hayenga, A. O., & Henderlong Corpus, J. (2010). Profiles of intrinsic and extrinsic motivations: A person-centered approach to motivation and achievement in middle school. *Motivation and Emotion, 34*, 371–383.

Haynes, T. L., Ruthig, J. C., Perry, R. P., Stupnisky, R. H., & Hall, N. C. (2006). Reducing the academic risks of over-optimism: The longitudinal effects of attributional retraining on cognition and achievement. *Research in Higher Education, 47*(7), 755–779.

Hidi, S. & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*, 111–127.

Hofer, B. K., & Yu, S. L. (2003). Teaching self-regulated learning through a" Learning to Learn" course. *Teaching of Psychology, 30*(1), 30–33

Hong, Y. Y., Chiu, C. Y., Dweck, C. S., Lin, D. M. S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology, 77*(3), 588–599

Hoyert, M. S., & O'Dell, C. D. (2006). A brief intervention to aid struggling students: A case of too much motivation? *Journal of Scholarship of Teaching and Learning, 6*(1), 1–13.

Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology, 49*(5), 505–528.

Hudley, C., Graham, S., & Taylor, A. (2007). Reducing aggressive behavior and increasing motivation in school: The evolution of an intervention to strengthen school adjustment. *Educational Psychologist, 42*, 251–260.

Hulleman, C. S., An, B., Hendricks, B., & Harackiewicz, J. M. (2007, June). *Interest development, achievement, and continuing motivation: The pivotal role of utility value*. Poster presented at the Institute of Education Sciences Research Conference, Washington, DC.

Hulleman, C. S., Durik, A. M., Schweigert, S., & Harackiewicz, J. M. (2008). Task values, achievement goals, and interest: An integrative analysis. *Journal of Educational Psychology, 100*, 398–416.

Hulleman, C. S., Godes, O., Hendricks, B., & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology, 102*(4), 880–895.

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326*, 1410–1412.

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development, 73*, 509–527.

Jamieson, J. P., Mendes, W. B., Blackstock, E., & Schmader, T. (2010). Turning the knots in your stomach into bows: Reappraising arousal improves performance on the GRE. *Journal of Experimental Social Psychology, 46*(1), 208–212.

Johnson, M. L., & Sinatra, G. M. (2013). Use of task-value instructional inductions for facilitating engagement and conceptual change. *Contemporary Educational Psychology, 38*, 51–63.

Juvonen, J. (2007). Reforming middle schools: Focus on continuity, social connectedness, and engagement. *Educational Psychologist, 42*, 197–208.

Kitsantas, A., Robert, A. R., & Doster, J. (2004). Developing self-regulated learners: Goal setting, self-evaluation, and organizational signals during acquisition of procedural skills. *The Journal of Experimental Education, 72*, 269–287.

Kosovich, J. J., Hulleman, C. S., Barron, K. E., & Getty, S. (2015). Developing a practical measure of motivation: Expectancy-Value-Cost in middle school science and mathematics. *Journal of Early Adolescence, 35*(5-6), 790–816.

Lavender, J. S. (2005). *The relationship between locus of control and academic success in college*. *Dissertation Abstracts International, 67*(1), 60.

Lazowski, R. A., & Hulleman, C. S. (2013, May). *Motivation interventions in education: A Meta-analytic review*. Poster presented at the annual meeting of the American Educational Research Association, San Francisco.

Lazowski, R. A., & Hulleman, C. S. (2015). Motivation interventions in education: A meta-analysis. *Review of Educational Research*. Online first edition.

Lefcourt, H. M., Von Baeyer, C. L., Ware, E. E., & Cox, D. J. (1979). The multidimensional multiattributional causality scale: The development of a goal specific locus of control scale. *Canadian Journal of Behavioural Science/Revue Canadienne des Sciences du Comportement, 11*(4), 286–304.

Lepper, M. R., & Henderlong, J. (2000). Turning "play" into "work" and "work" into "play": 25 years of research on intrinsic versus extrinsic motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 257–307). New York: Academic.

Lepper, M. R., Henderlong, J., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *Journal of Educational Psychology, 97*, 184–196.

Lewin, K., Dembo, T., Festinger, L., & Sears, P. S. (1944). Level of aspiration. In J. M. Hunt (Ed.), *Personality and behavior disorders* (pp. 333–378). New York: The Ronald Press.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Englewood Cliffs, NJ: Prentice Hall.

Marinak, B. A., & Gambrell, L. B. (2008). Intrinsic motivation and rewards: What sustains young children's engagement with text? *Literacy Research and Instruction, 47*, 9–26.

Marsh, H. W. (1990). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review, 2*(2), 77–172.

Marsh, H. W. (1992). *Self Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescent self-concept. A test manual and research monograph*. South Penrith, NSW: University of Western Sydney, Faculty of Education.

Marsh, H. W., Byrne, B. M., & Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement: Reanalysis of a pioneering study and revised recommendations. *Educational Psychologist, 34*(3), 155–167.

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology, 81*(1), 59–77.

Marsh, H. W., & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist, 20*(3), 107–123.

Martin, A. J. (2008). Enhancing student motivation and engagement: The effects of a multidimensional intervention. *Contemporary Educational Psychology, 33*(2), 239–269.

Miller, S. D., & Meece, J. L. (1997). Enhancing elementary students' motivation to read and write: A classroom intervention study. *The Journal of Educational Research, 90*(5), 286–299.

Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science, 330*(6008), 1234–1237.

Morisano, D., Hirsh, J. B., Peterson, J. B., Pihl, R. O., & Shore, B. M. (2010). Setting, elaborating, and reflecting on personal goals improves academic performance. *Journal of Applied Psychology, 95*, 255–264.

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology, 75*(1), 33–52.

Muis, K. R., Ranellucci, J., Franco, G. M., & Crippen, K. J. (2013). The interactive effects of personal achievement goals and performance feedback in an undergraduate science class. *The Journal of Experimental Education, 81*(4), 556–578.

Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K. T., & Trautwein, U. (2011). Who took the "x" out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science, 22*(8), 1058–1066.

Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education, 7*(2), 133–144.

Oyserman, D., Terry, K., & Bybee, D. (2002). A possible selves intervention to enhance school involvement. *Journal of Adolescence, 25*, 313–326.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543–578.

Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory into Practice, 41*(2), 116–125.

Patall, E. A., Cooper, H., & Wynn, S. R. (2010). The effectiveness and relative importance of choice in the classroom. *Journal of Educational Psychology, 102*(4), 896–915.

Patrick, B. C., Hisley, J., & Kempler, T. (2000). "What's everybody so excited about?": The effects of teacher enthusiasm on student intrinsic motivation and vitality. *The Journal of Experimental Education, 68*(3), 217–236.

Patrick, B. C., Skinner, E. A., & Connell, J. P. (1993). What motivates children's behavior and emotion? Joint effects of perceived control and autonomy in the academic domain. *Journal of Personality and Social Psychology, 65*(4), 781–791.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315–341.

Perez, T., Cromley, J. G., & Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. *Journal of Educational Psychology, 106*(1), 315–329.

Perry, R. P., Stupnisky, R. H., Hall, N. C., Chipperfield, J. G., & Weiner, B. (2010). Bad starts and better finishes: Attributional retraining and initial performance in competitive achievement settings. *Journal of Social and Clinical Psychology, 29*(6), 668–700.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology, 95*(4), 667–686.

Ramirez, G., & Beilock, S. L. (2011). Writing about testing worries boosts exam performance in the classroom. *Science, 331*(6014), 211–213.

Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and Emotion, 28*, 147–169.

Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist, 46*(3), 168–184.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*, 353–387.

Rimm-Kaufman, S. E., & Hulleman, C. S. (2015). Social and emotional learning in elementary school settings: Identifying mechanisms that matter. In: J. P. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.), *The Handbook of social and emotional learning: Research and practice* (pp. 151-166). Guilford Press.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*(2), 261–288.

Rokeach, M. (1973). *The nature of human values*. New York: Free Press.

Romero, C., Master, A., Paunesku, D., Dweck, C. S., & Gross, J. J. (2014). Academic and emotional functioning in middle school: The role of implicit theories. *Emotion, 14*(2), 227–234.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied, 80*(1), 1–28.

Ruthig, J. C., Perry, R. P., Hall, N. C., & Hladkyj, S. (2004). Optimism and attributional retraining: Longitudinal effects on academic achievement, test anxiety, and voluntary course withdrawal in college students. *Journal of Applied Social Psychology, 34*(4), 709–730.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78.
Sansone, C., & Harackiewicz, J. M. (Eds.). (2000). *Intrinsic and extrinsic motivation: The search for optimal motivation and performance*. San Diego, CA: Academic.
Schaffner, E., & Schiefele, U. (2007). The effect of experimental manipulation of student motivation on the situational representation of text. *Learning and Instruction, 17*, 755–772.
Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26*(3–4), 207–231.
Schunk, D. H., & Pajares, F. (2002). The development of academic self-efficacy. In A. Wigfield & J. Eccles (Eds.), *Development of achievement motivation* (pp. 15–31). San Diego, CA: Academic.
Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values? *Journal of Social Issues, 50*(4), 19–45.
Schwartz, S. H., & Bilsky, W. (1990). Toward a theory of the universal content and structure of values: Extensions and cross-cultural replications. *Journal of Personality and Social Psychology, 58*(5), 878.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin Company.
Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*(3), 407–441.
Sheldon, K. M., Elliot, A. J., Kim, Y., & Kasser, T. (2001). What is satisfying about satisfying events? Testing 10 candidate psychological needs. *Journal of Personality and Social Psychology, 80*(2), 325–339. doi:10.1037/0022-3514.80.2.325.
Shell, D. F., Murphy, C. C., & Bruning, R. H. (1989). Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. *Journal of Educational Psychology, 81*(1), 91–100. doi:10.1037/0022-0663.81.1.91.
Sherman, D. K., Hartson, K. A., Binning, K. R., Purdie-Vaughns, V., Garcia, J., Taborsky-Barba, S., et al. (2013). Deflecting the trajectory and changing the narrative: How self-affirmation affects academic performance and motivation under identity threat. *Journal of Personality and Social Psychology, 104*, 591–618.
Silva, J. P., White, G. P., & Yoshida, R. K. (2011). The direct effects of principal–student discussions on eighth grade students' gains in reading achievement: An experimental study. *Educational Administration Quarterly, 47*(5), 772–793.
Simons, J., Dewitte, S., & Lens, W. (2003). "Don't do it for me. Do it for yourself!" Stressing the personal relevance enhances motivation in physical education. *Journal of Sport and Exercise Psychology, 25*, 145–160.
Skaalvik, E. M., & Rankin, R. J. (1995). A test of the internal/external frame of reference model at different levels of math and verbal self-perception. *American Educational Research Journal, 32*(1), 161–184.
Skaalvik, E. M., & Rankin, R. J. (1996). *Self-concept and self-efficacy: Conceptual analysis*. Paper presented at the annual meeting of the American Educational Research Association, New York.
Skinner, E. A. (1995). *Perceived control, motivation, and coping*. Thousand Oaks, CA: Sage.
Skinner, E. A. (1996). A guide to constructs of control. *Journal of Personality and Social Psychology, 71*, 549–570.
Skinner, E. A., Wellborn, J. G., & Connell, J. P. (1990). What it takes to do well in school and whether I've got it: A process model of perceived control and children's engagement and achievement in school. *Journal of Educational Psychology, 82*(1), 22–32.
Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613–629.
Struthers, C. W., & Perry, R. P. (1996). Attributional style, attributional retraining, and inoculation against motivational deficits. *Social Psychology of Education, 1*(2), 171–187.
Stupnisky, R. H., Perry, R. P., Hall, N. C., & Guay, F. (2012). Examining perceived control level and instability as predictors of first-year college students' academic achievement. *Contemporary Educational Psychology, 37*(2), 81–90.

Tang, S., & Hall, V. C. (1995). The overjustification effect: A meta-analysis. *Applied Cognitive Psychology, 9*, 365–404.

Tighezza, M. (2013). Modeling relationships among learning, attitude, self-perception, and science achievement for grade 8 Saudi students. *International Journal of Science and Mathematics Education, 12*(4), 721–740.

Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Journal of Educational Psychology, 104*(3), 763–777.

Tunnell, G. B. (1977). Three dimensions of naturalness: An expanded definition of field research. *Psychological Bulletin, 84*, 426–437.

Updegraff, K. A., Eccles, J. S., Barber, B. L., & O'Brien, K. M. (1996). Course enrollment as self-regulatory behavior: Who takes optional high school math courses? *Learning and Individual Differences, 8*, 239–259.

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*(4), 751–796.

Vallerand, R. J., & Blssonnette, R. (1992). Intrinsic, extrinsic, and amotivational styles as predictors of behavior: A prospective study. *Journal of Personality, 60*(3), 599–620.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Brière, N. M., Senecal, C., & Vallieres, E. F. (1993). On the assessment of intrinsic, extrinsic, and amotivation in education: Evidence on the concurrent and construct validity of the Academic Motivation Scale. *Educational and Psychological Measurement, 53*(1), 159–172.

Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K. M., & Deci, E. L. (2004). Motivating learning, performance, and persistence: The synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *Journal of Personality and Social Psychology, 87*(2), 246–260.

Vansteenkiste, M., Simons, J., Lens, W., Soenens, B., & Matos, L. (2005). Examining the motivational impact of intrinsic versus extrinsic goal framing and autonomy-supportive versus internally controlling communication style on early adolescents' academic achievement. *Child Development, 76*, 483–501.

Vansteenkiste, M., Timmermans, T., Lens, W., Soenens, B., & Van den Broeck, A. (2008). Does extrinsic goal framing enhance extrinsic goal-oriented individuals' learning and performance? An experimental test of the match perspective versus self-determination theory. *Journal of Educational Psychology, 100*(2), 387–397.

Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology, 92*(1), 82–96.

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes among minority students. *Science, 331*, 1447–1451.

Weiner, B. (1972). *Theories of motivation: From mechanism to cognition*. Oxford, England: Markham.

Weiner, B. (1980). The role of affect in rational (attributional) approaches to human motivation. *Educational Researcher, 9*(7), 4–11.

Weiner, B. (1992). *Human motivation: Metaphors, theories, and research*. Thousand Oaks, CA: Sage.

Weiner, B. (2010). The development of an attribution-based theory of motivation: A history of ideas. *Educational Psychologist, 45*(1), 28–36.

Weiner, B., Nierenberg, R., & Goldstein, M. (1976). Social learning (locus of control) versus attributional (causal stability) interpretations of expectancy of success1. *Journal of Personality, 44*(1), 52–68.

Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review, 30*(1), 1–35.

Wigfield, A., & Eccles, J. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review, 12*, 265–310.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of motivation. *Contemporary Educational Psychology, 25*, 68–81.

Wigfield, A., & Eccles, J. (2002). *Development of achievement motivation*. San Diego, CA: Academic.

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A., Freedman-Doan, C., et al. (1997). Changes in children's competence beliefs and subjective task values across the elementary school years: A three-year study. *Journal of Educational Psychology, 89*, 451–469.

Wigfield, A., & Wentzel, K. R. (2007). Introduction to motivation at school: Interventions that work. *Educational Psychologist, 42*(4), 261–271.

Wilson, T. D., & Linville, P. W. (1982). Improving the academic performance of college freshmen: Attribution therapy revisited. *Journal of Personality and Social Psychology, 42*(2), 367.

Wilson, T. D., & Linville, P. W. (1985). Improving the performance of college freshmen with attributional techniques. *Journal of Personality and Social Psychology, 49*(1), 287.

Xiang, P., Chen, A., & Bruene, A. (2005). Interactive impact of intrinsic motivators and extrinsic rewards on behavior and motivation outcomes. *Journal of Teaching in Physical Education, 24*, 179–197.

Yeager, D. S., & Bundick, M. J. (2009). The role of purposeful work goals in promoting meaning in life and in schoolwork during adolescence. *Journal of Adolescent Research, 24*, 423–452.

Yeager, D. S., Paunesku, D., Walton, G. M., & Dweck, C. S. (2013). *How can we instill productive mindsets at scale? A review of the evidence and an initial R&D agenda*. A white paper prepared for the White House meeting on "Excellence in Education: The Importance of Academic Mindsets." Austin, TX: The University of Texas.

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*(2), 267–301.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29*, 663–676.

Chapter 11
Academic Emotions and Their Regulation via Emotional Intelligence

Thomas Goetz and Madeleine Bieg

11.1 Introduction

"I felt angry because I didn't contribute anything to the lesson, and I wasn't able to control that anger." This statement from a 15-year-old student, who participated in a diary study about emotions in the classroom, illustrates the salience of academic emotions and the regulation of academic emotion in everyday classroom experiences. Students possess knowledge about the significance of controlling some emotions in order to be able to learn and achieve yet may not always be capable of doing so. Generally speaking, emotions are important to well-being and are related to school achievement. Thus, regulating one's emotions in a way that is beneficial for learning and achievement is an important noncognitive skill to focus on in the context of education.

The present chapter explores academic emotions and the role emotional intelligence plays in emotion regulation. We will elucidate a number of intervention strategies that can be used to foster students' emotional intelligence and help them become experts of their own emotions. Emotional intelligence is multifaceted and includes the capacity to recognize and reflect on one's emotions, general knowledge of emotions, and an understanding of which emotions are conducive to learning. Successfully regulating one's emotions, however, may make the biggest difference

T. Goetz (✉)
Empirical Educational Research, University of Konstanz, Universitätsstraße 10, 78464 Konstanz, Germany

Thurgau University of Teacher Education, Unterer Schulweg 3, Postfach, 8280 Kreuzlingen, Switzerland
e-mail: thomas.goetz@uni-konstanz.de

M. Bieg
Empirical Educational Research, University of Konstanz, Universitätsstraße 10, 78464 Konstanz, Germany
e-mail: madeleine.bieg@uni-konstanz.de

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_11

in terms of learning and achievement. For example, the student mentioned above would likely have benefitted from using better self-regulatory strategies, particularly the ability to regulate the emotion of anger. Aside from gaining general knowledge about emotions, enhancing students' capacity to regulate their emotions is a worthwhile aim to accomplish in the course of schooling. This chapter intends to promote greater awareness among researchers and practitioners of the critical role that academic emotions play for academic attainment, emphasizing how teaching students the importance of academic emotions might enhance learning and achievement.

11.2 Defining and Classifying Academic Emotions

Various conceptions of emotions exist in the research literature. In the present chapter, we focus on academic emotions, or emotions that are specifically related to learning and achievement. It is worth noting that some definitions presented in the following section apply to (or even stem from) definitions of emotions in general, but apply to academic emotions as well. Typical emotions experienced in the context of school include pride, enjoyment, hope, anxiety, anger, shame, and, most frequently, boredom (e.g., Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011).

Emotions are distinguished from other related constructs that are important to learning and achievement, such as mood or well-being. One the one hand, mood differentiates itself from emotions by connoting a less intense, enduring psychological state that is not related to a specific cause (Frenzel & Stephens, 2013). Well-being, on the other hand, involves the satisfaction of needs and the experience of positive emotions more often than negative emotions. Well-being has its own tradition in psychological research (Frenzel & Stephens, 2013).

Emotions are defined as multifaceted constructs comprising five distinct components (Frenzel & Stephens, 2013; building on the original research by Dodge, 1989). These are: the affective component (i.e., the core "feeling"), the cognitive component (i.e., thoughts), the motivational component (i.e., motivation for a behavior), the physiological component (e.g., arousal, heart rate, etc.), and the expressive component (e.g., facial expression, gestures). For example, the emotion of enjoyment might include the experience of a "happy" feeling, thoughts about how fun an activity is, motivation to continue the activity, increased heart rate, and the expression of a smile on the face.

Emotions in general can be classified according to valence (positive versus negative) and level of arousal (activating versus deactivating; Russell, 1980). In the case of academic emotions, Pekrun, Elliot, and Maier (2006) propose a classification according to valence (positive versus negative), object focus (activity versus outcome), and temporal focus (prospective, retrospective, or current). Pride, for example, is a positive, activating emotion that is retrospectively focused on the outcome of an activity related to learning or achievement. Anger, by contrast, is a negative, activating emotion related to activities with a current temporal focus, or to outcomes with a retrospective temporal focus.

Further specifications apply to particular academic emotions experienced in school settings. For the emotion of anxiety experienced specifically in a testing situation (test anxiety), for example, the worry component and the emotionality component are delineated (e.g., Zeidner, 1998). The worry component includes a cognitive aspect, such as thoughts about possible failure, while the emotionality component includes affective, physiological, and expressive facets. Another academic emotion is boredom, which can be classified into different types according to levels of arousal and valence. Five boredom types are characterized by different levels of valence and arousal: indifferent boredom, calibrating boredom, searching boredom, reactant boredom, and apathetic boredom (Goetz et al., 2014).

Distinctions are also made between trait and state emotions. Trait emotions are enduring individual tendencies to experience a specific emotion in identical or similar situations (Amelang, Bartussek, Stemmler, & Hagemann, 2006), or repeatedly occurring emotional states (i.e., habitual emotions). State emotions, by contrast, are more unstable and transient emotional experiences (Eid, Schneider, & Schwenkmezger, 1999). In an academic setting, a student's experience of anxiety might reflect a stable characteristic of the student (i.e., he or she is an anxious person), or be tied to a specific situation at school (e.g., anxious because of a forthcoming exam). The distinction between traits and states is also relevant to the assessment of emotions, which is further elaborated in the following section of this chapter.

11.3 Assessing Academic Emotions

Researchers employ a number of approaches to assess emotions. These may include physiological, observational, or self-report measures designed to capture intensity and frequency of occurrence of specific emotions. While most of the assessment methods explained in this section were not originally developed to measure academic emotions, they can easily be adapted for this purpose.

Physiological measures of emotions include skin conductance, heart rate, and cortisol. Imaging techniques, such as functional magnetic resonance imaging (fMRI) or electroencephalography (EEG), are also used to measure emotions. However, a major disadvantage of these techniques is that identifying discrete emotions is hardly possible. Another commonly used method is deducing emotions from facial expressions (e.g., Reisenzein, Junge, Studtmann, & Huber, 2014). One prominent observational measure of this type is Ekman's "Facial Action Coding System," which has been transferred into a computer program for facial detection (FACS; e.g., Cohn, Ambadar, & Ekman, 2007; Ekman, Friesen, & Hager, 2002; Ekman & Rosenberg, 2005). Another recently developed computer program for identifying and measuring emotions via facial expressions is the FaceReader™ (e.g., D'Arcey, Johnson, & Ennis, 2012; Noldus Information Technology, 2012). Observational measures, however, are time consuming to gather, especially when assessing the emotions of many students in one classroom is the goal. For this reason, researchers in education typically rely on students' self-reports.

One classic self-report measure is the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988), which asks participants to rate the intensity of their feelings in reference to positive and negative adjectives. However, an ongoing discussion in the field concerns whether the assessment of the two poles of positive and negative emotion is sufficient (e.g., Barrett, 1998; Vansteelandt, van Mechelen, & Nezlek, 2005). An increasing number of researchers are choosing to assess discrete emotions, instead of simply distinguishing between positive and negative valence. This is because positive or negative emotions might emerge in different situations and relate differently to antecedent and outcome variables. For example, boredom and anxiety are two commonly experienced achievement emotions, and both are negatively related to academic self-concept. However, this relationship is much stronger for anxiety (Goetz, Cronjaeger, Frenzel, Lüdtke, & Hall, 2010). Besides the PANAS, several self-report instruments have been developed to assess emotions. One emotion-specific questionnaire used to measure anxiety is the State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970), which assesses anxiety as moment-specific state and also as a stable trait. The Academic Emotions Questionnaire (AEQ; Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011) assesses a wide range of academic-specific emotions and is particularly relevant for use in learning and achievement contexts. An alternative to heavily text-based questionnaires that is suitable for younger children is the Self-Assessment Manikin (Lang, 1980). This instrument uses pictures of small cartoon figures to assess dimensions such as arousal, valence, and dominance. Table 11.1 displays an overview of several

Table 11.1 Methods to assess emotions

Instrument	Aim	Advantages	Disadvantages
Physiological measures	Assessment of arousal, skin conductance, heart rate	Objective	Not able to assess discrete emotions
Imaging techniques (fMRI, EEG)	Identify active brain regions	Objective	Labor intensive, not able to assess discrete emotions
Facial Action Coding System (FACS)	Assessment of emotions by observation	Relatively objective	Difficult to learn
Self-Assessment Manikin (SAM)	Assessment of arousal, valence, and dominance	Usable with younger children	Not able to assess discrete emotions
Positive Affect Negative Affect Scale (PANAS)	Assessment of positive and negative affect by the use of adjective scales	Easy to apply, comprehensive assessment	Only poles of positive and negative affect, relatively subjective, eventually biased
Academic Emotions Questionnaire (AEQ)	Assessment of positive and negative achievement emotions (class, test, learning)	Comprehensive assessment of academic emotions in a wide range of situations	Relatively subjective, eventually biased

assessment methods (not only for academic emotions) and describes possible advantages and disadvantages of each method.

When deciding on how to assess emotions in the context of schooling and achievement, it is important to consider whether emotions will be measured in a single academic domain or across academic domains. However, because previous research indicates that academic emotions are domain-specific (e.g., Goetz, Pekrun, Hall, & Haag, 2006), a domain-specific assessment may be more suitable (similar to the domain-specific assessment of self-concept; Brunner et al., 2010).

Several possible methods exist for assessing academic emotions and deciding which one to choose may be situation dependent. However, it is worth highlighting a recent development in how trait and state emotions are measured via self-report. As mentioned previously, researchers in the field of emotion assessment distinguish between generally experienced emotions (traits or habitual emotions) and emotions experienced in real time (state emotions). This distinction is not new, as the two types of emotions were already operationalized in the STAI (Spielberger et al., 1970). Measuring state emotions using the experience-sampling method provides researchers with information about the fluctuation of emotions as they occur in "real time" (Csikszentmihalyi & Larson, 1987). This differs greatly from the more convenient and static assessment of trait emotions. An important point to note, however, is that empirical findings and theoretical assumptions suggest trait emotions are strongly influenced by semantic knowledge and beliefs, whereas state emotions are less prone to memory biases based on a person's belief system (Pekrun & Bühner, 2014; Robinson & Clore, 2002). Trait emotions, therefore, may rely on beliefs about emotions more than state emotions do, which are assessed more directly. According to Robinson and Clore (2002), episodic memory is used with state emotional assessments, whereas semantic memory is used when general emotions are assessed.

In line with this differentiation between trait and state emotions, one recent study on mathematics anxiety in boys and girls found gender differences in students' trait math anxiety, but not in state math anxiety (Goetz, Bieg, Lüdtke, Pekrun, & Hall, 2013). The higher discrepancy between trait and state anxiety in girls as compared to boys was due to their lower competence beliefs in the domain of mathematics. This finding illustrates how the type of emotional assessment a researcher chooses might lead to very different findings. Thus, the choice of assessment should always be based on the research question at hand.

11.4 Relationship Between Academic Emotions and Key Constructs

Aside from being directly related to learning and achievement, emotions are relevant outcomes in and of themselves and are associated with several key constructs that are important in academic contexts. For example, academic emotions are related to antecedent variables such as appraisal of control and value. Understanding

these antecedents may provide greater knowledge about how to influence emotions in a way that is more conducive to learning and achievement.

11.4.1 Theoretical Model: Control-Value Theory of Achievement Emotions

Appraisal theories are particularly relevant when considering antecedents of emotions. Specifically, Pekrun's (2006) control-value theory is a prominent approach to understanding precursors to academic emotions, as well as the relationship between emotions and academic outcomes. This theoretical model holds that the cognitive appraisal antecedents of control and value are central to the emergence of several emotions. Pekrun's theory describes social and environmental factors that influence appraisal antecedents and also specifies the link between emotions and achievement via mediating variables (Pekrun & Perry, 2014). The model is depicted in Fig. 11.1 and will be elaborated on in the following sections, beginning with a general discussion of the relationship between academic emotions and achievement and then moving to examine specific antecedents of academic emotions.

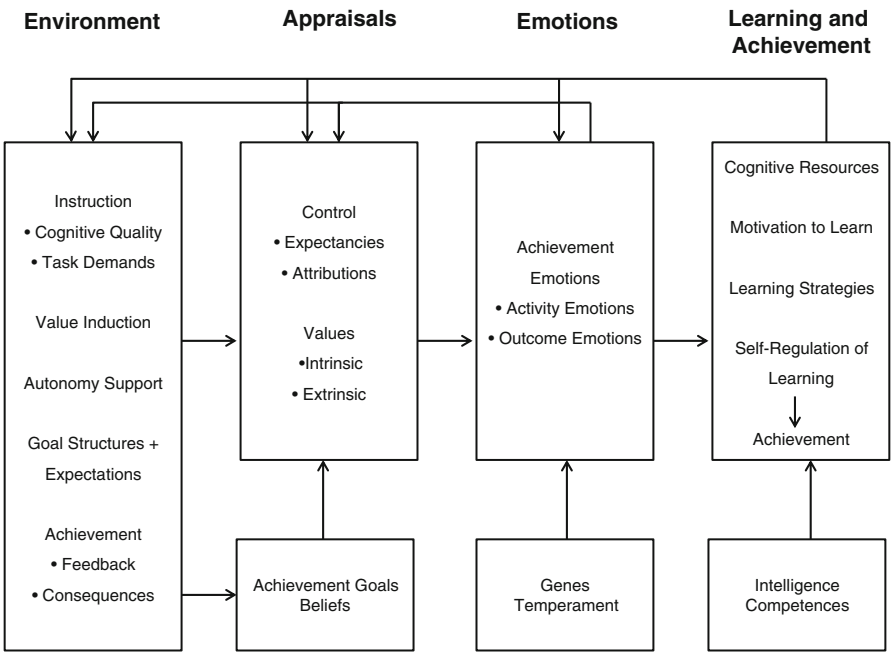

Fig. 11.1 The control-value theory of achievement emotions (adapted from Pekrun (2006), with permission of Springer Science + Business Media)

11.4.2 Academic Emotions and Achievement

Aside from intelligence and other cognitive characteristics, emotions play a significant role in the learning process. According to a recently published review, the overall correlation between emotions and achievement (based on results from single studies) is $r = \left| 0.25 \right|$ (Goetz & Hall, 2013). In particular, the relationship between test anxiety and achievement is well studied, and meta-analyses report negative effect sizes ranging from −0.20 to −0.25 (Goetz & Hall, 2013). Generally, positive emotions show a positive relationship to achievement, whereas negative emotions relate negatively to achievement. Control-value theory (Pekrun, 2006) posits that the relationship between emotions and achievement is mediated by variables such as cognitive resources, motivation to learn, use of learning strategies, and self-regulation of learning. Moderators of the emotion-achievement relationship have been demonstrated empirically. One moderator is subject domain, where a stronger relationship between emotions and achievement has been found in math and science as compared to verbal domains. Emotion valence is also a determinant of whether the relationship between emotions and learning outcomes is positive or negative (Goetz & Hall, 2013). With anxiety, it has also been found that the worry component exerts a more detrimental influence on learning and achievement because it is present during the entire testing situation, whereas the emotionality component attenuates over time (Galassi, Frierson, & Sharer, 1981).

As mentioned previously, it is assumed the path between emotions and achievement is mediated by a number of variables. All emotions, positive and negative, draw on cognitive resources that may subsequently become unavailable for solving cognitive tasks. This might lead to the conclusion that a relatively neutral mood is the most desirable state for enhancing achievement, especially if emotions are not directly related to the learning or academic activity (Frenzel & Stephens, 2013; Meinhardt & Pekrun, 2003). However, recent findings indicate that positive emotions usually enhance motivation to learn, whereas negative emotions – at least in the long run – are negatively correlated with motivation to learn (Frenzel & Stephens, 2013). Over the short term, however, negative emotions such as test anxiety may externally motivate learning (Pekrun, 1992) and can lead to better achievement outcomes. In terms of learning strategies, positive emotions promote the use of self-regulated learning and favorably influence achievement. Learners experiencing negative emotions, by contrast, rely more heavily on external guides and use more rigid learning strategies (Frenzel & Stephens, 2013). It is worth noting that feedback loops may stem from mediating variables as well as academic emotions themselves. For example, effective self-regulation might enhance positive emotions, and academic success might relate to positive (outcome) emotions. However, empirical findings regarding potential feedback processes are sparse and further investigation is needed. In sum, there is a significant relationship between emotions and achievement, which is influenced by a number of mediating and moderating variables. Knowledge of these variables, we argue, offers possibilities about how to influence the relationship between emotion and achievement.

11.4.3 Antecedents of Academic Emotions

Academic emotions are influenced by learners' subjective cognitive appraisals, in addition to other factors such as temperament or genetics. Given that the same person may experience different emotions in different situations and that different people may experience different emotions in the same situation, appraisal theories of emotions are the most commonly considered theoretical approach to (academic) emotions. The subjective appraisals of control and value are particularly salient antecedents of academic emotions. Control-value theory (Pekrun, 2006) posits that level of control is positively related to positive emotions and negatively related to negative emotions. Value is the subjective importance of an activity or outcome and is positively related to all emotions such that high importance placed on an activity or outcome intensifies all emotions except boredom. Beyond individual main effects, theoretical assumptions and empirical findings suggest that control and value appraisals interact in predicting emotions (e.g., Bieg, Goetz, & Hubbard, 2013), meaning the relationship between an emotion and the respective level of an appraisal (i.e., control) is dependent on the extent of the other appraisal (i.e., value). For example, the effect of control on anxiety should be stronger in situations where the outcome is very important (high value), as compared to situations where less value is placed on the outcome. Situation-specific appraisals may also be influenced by achievement goals and beliefs.

According to the control-value model, certain social and environmental variables may influence appraisals and therefore academic emotions. These variables are instruction quality, value induction, autonomy support, goal structures and expectations, feedback on achievement, and possible consequences of achievement. Delivering clear and understandable instructions to students, for example, can positively influence students' competence beliefs. Students' competence beliefs also increase when they are permitted autonomy and recognize their capacity to master difficult tasks. In addition, classroom goal structures are closely tied to learners' competence appraisals. A focus on mastering learning material and working through challenging tasks is more emotionally advantageous than an explicit focus on competition. Expectations of students also influence their emotions. If a teacher, for example, expects a student to succeed or not to succeed, the student's control appraisal (and emotional response) will follow accordingly. Control beliefs are also modified by providing feedback to students about their level of competence or mastery. Emphasizing the importance of learning and performance outcomes might also enhance students' perception of an academic task as valuable.

Similar to the emotion-achievement relationship described previously, feedback loops also exist between emotions and their antecedents. Students who display different emotions might provoke different responses from their environment. For example, a helpless student might evoke more helping behavior from his or her environment, which in turn may strengthen the student's feelings of helplessness. By contrast, a student showing positive emotions about learning might provoke

more interaction with a teacher, which in turn may result in the student's experience of more positive emotions.

The control-value approach detailed previously is only one possible explanation for the emergence of emotions, among others (e.g., evolutionary or social theories). However, few theories explicitly focus on the development of achievement emotions. In addition to control-value theory, attribution theory has a long tradition in learning and psychological research and has been utilized to understand the emergence of academic emotions (Weiner, 1985). Attribution theory is concerned with the assumptions a person makes about the causes of their success or failure, and how this attribution affects subsequent outcomes of interest. Three dimensions are distinguished in this approach, namely, how stable or unstable the cause is, whether it is internal or external, and whether it is controllable or uncontrollable (Graham & Taylor, 2014). It is important to identify the causes a student attributes to their success or failure to understand how attributions influence the emergence of emotions. A student attributing his or her failure to stable and internal causes (such as low ability) may experience heightened negative emotions. Attributing success to controllable and internal causes, in turn, is more strongly related to positive emotions and perseverance, particularly when students encounter failure (Dweck, 2002).

11.5 Emotional Intelligence as a Central Approach for Regulating Academic Emotions

As outlined above, students' academic emotions play an important role in learning and achievement (see also Garner, 2010). Identifying ways to foster academic emotions that are beneficial to students' learning processes and achievement outcomes is an important area of research. Facilitating students' development of beneficial academic emotions can be based on interventions initiated by persons having a direct (or mediated) impact on students' academic emotional experiences. For example, according to Pekrun's (2006) control-value theory, altering aspects of students' environment can enhance students' academic emotions (e.g., ways of teaching, parental involvement, peer attitudes about learning and achievement). With respect to lifelong learning (cf., Schober et al., 2007) and individualizing the learning processes, a more meaningful approach is to teach students how to independently regulate their emotions. Numerous strategies for regulating one's emotions are outlined in the literature on coping (e.g., Lazarus & Folkman, 1984; Rost & Schermer, 1987; Zeidner & Endler, 1996). Three basic types of coping are typically differentiated: emotion focused, problem focused, and avoidance. Emotion-focused coping refers to a modification of one's own emotions, such as adopting relaxation techniques to reduce anxiety. Problem-focused coping refers to a modification of emotion-inducing circumstances, like restructuring the environment. Lastly, avoidance coping involves mentally or physically avoiding an emotion-eliciting situation.

Comprehensive approaches to regulating one's own emotions (and those of others) are outlined in the field of emotional intelligence (EI), which partly overlaps with methods of emotion regulation developed in the research on coping (cf., Zeidner, Matthews, & Roberts, 2012). Subskills contributing to successful emotion regulation discussed in detail in the field of EI include the ability to identify one's own emotions and to know about the effects of emotions – both aspects represent prerequisites for judging whether initiating a coping process is necessary at all. Given that EI approaches to emotion regulation provide a comprehensive framework and expand strategies outlined in the classic coping literature, we focus on EI as a keystone skill for regulating one's own academic emotions.

11.6 A Model for the Promotion of Emotional Intelligence in Learning and Achievement Situations: The PEILAS Model

EI has become a popular construct of interest in the past 20 years (see Mayer, Roberts, & Barsade, 2008; for a review detailing the history of the construct see Allen, MacCann, Matthews, & Roberts, 2014). Since the 1990s, numerous programs aimed at promoting EI in school settings have been developed, with many being published in the USA in the context of social and emotional learning (SEL; Brackett & Rivers, 2014; Cohen, 1999, 2001; Elias, Hunter, & Kress, 2001; Topping, Holmes, & Bremner, 2000). An example of a still very popular SEL program, which focuses on fostering EI, is the PATHS program (Promoting Alternative Thinking Strategies; Greenberg, Kusche, Cook, & Quamma, 1995; for a description of the program, see also Brackett & Rivers, 2014; Parker, Saklofske, Wood, & Collin, 2009). However, two central problems exist in most school-based programs designed to foster EI. First, they usually lack a clear definition of EI. Second, these programs typically refer to emotions related to social contexts, but not to emotions related to learning and achievement. Having taken these shortcomings into account, Goetz, Frenzel, Pekrun, and Hall (2005) developed a model for fostering EI in the context of learning and achievement. This model allows for numerous practical implications and can be used as a heuristic for designing EI-based intervention and promotion programs in school contexts. In this section, we outline a revised version of this framework – the PEILAS model (for the **P**romotion of **EI** in **L**earning and **A**chievement **S**ituations).

11.6.1 Defining Emotional Intelligence in the PEILAS Model

Various definitions and conceptualizations of "emotional intelligence" exist in the literature (see Matthews, Roberts, & Zeidner, 2004; Matthews, Zeidner, & Roberts, 2002). For the current PEILAS model, we remain consistent with the definition of

EI used to develop an earlier version of this framework. Goetz et al. (2005) outline a particularly suitable definition of EI for application in learning and achievement situations based on the Revised Ability Model of Emotional Intelligence by Mayer and Salovey (1997; see also Salovey & Mayer, 1990). Mayer and Salovey define EI as a mental ability construct and integrate four facets (or "branches") of emotional abilities (i.e., skills), outlined as follows: branch I, perception, appraisal, and expression of emotion; branch II, emotional facilitation of thinking; branch III, understanding and analyzing emotions and employing emotional knowledge; and branch IV, reflective regulation of emotions to promote emotional and intellectual growth. This approach to EI is (1) consistent with a cognitive conceptualization of intelligence (ability model of EI, see Allen et al., 2014; Brackett, Lopes, Ivcevic, Mayer, & Salovey, 2004; for contrasting "emotional intelligence" with "emotional competence" and "emotional literacy", see Humphrey, Curran, Morris, Farrell, & Woods, 2007; Lau & Wu, 2012), (2) requires a minimal number of school-specific modifications, (3) is suitable for operationalization and evaluation, and (4) is conducive to the development of intervention programs.

The PEILAS model incorporates facets of Mayer and Salovey's (1997) branches I, III, and IV. With respect to branch I, the PEILAS framework focuses on the perception of emotions; concerning branch III, it refers to reflection and knowledge about emotions (e.g., knowledge about its causes, emotional manifestations and effects, and methods of regulation); and for branch IV, the PEILAS model focuses on the management of one's own emotions (i.e., the capacity to regulate them). Branch II, which refers to generating and using emotions in an effective way, is less important because the PEILAS model mainly focuses on EI as regulating academic emotions, not as actively inducing or preventing specific emotions (for antecedent-focused vs. response-focused emotion regulation, see Gross, 1998, 2007). Furthermore, recent research indicates that the second branch cannot be conceptually or empirically distinguished from the other three branches (Joseph & Newman, 2010; MacCann, Joseph, Newman, & Roberts, 2014). To summarize, EI is defined in the PEILAS model as a person's cognitive ability to perceive, reflect, and regulate emotions. With respect to an evaluation of the PEILAS model, all three aspects (perception, reflection, regulation) can be assessed (e.g., by using the MSCEIT [Mayer Salovey Caruso Emotional Intelligence Test], as cited in Mayer, Salovey, & Caruso, 2002; for a youth version of the MSCEIT see Rivers et al., 2012).

11.6.2 Basic Assumptions of the PEILAS Model

The main theoretical assumptions underlying the current PEILAS model are similar to those outlined in the first version developed by Goetz et al. (2005). However, the revised version focuses more explicitly on the interactive effects between key variables in the model (see Fig. 11.2).

The PEILAS model incorporates approaches from the motivation research tradition in psychology, particularly facets of expectancy-value theory (Atkinson, 1957,

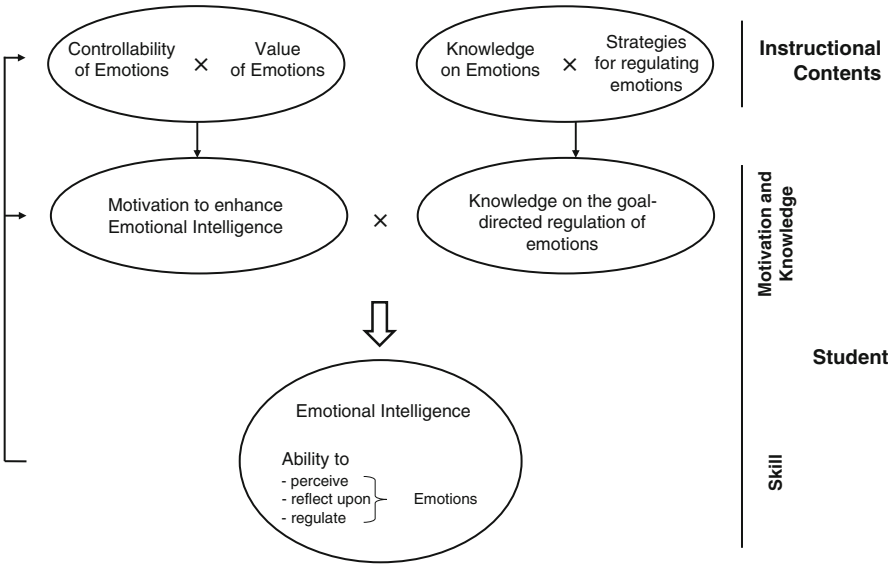

Fig. 11.2 The PEILAS model – a model for the promotion of EI in learning and achievement situations

1964). In this tradition, it is argued that motivation is based on (1) subjective appraisals of the probability of action outcomes, (2) the value of these outcomes to the individual, and (3) the interaction of both these aspects. Concerning the last point, it is assumed that both probability appraisals and value attributions must reach a minimum level for an individual to become motivated. Interestingly, this interaction effect has largely been ignored within research in the field of expectancy-value theory (see Bieg et al., 2013; Goetz, Frenzel, Stoeger, & Hall, 2010; Nagengast et al., 2011). Based on this observation, the PEILAS model holds that teaching *both* controllability (as a central aspect for enhancing appraisals of the probability of an action outcome) *and* value of emotions (see "Instructional Content" part of the model) is a necessary antecedent to students' motivation to enhance EI (see "Student" – "Motivation and Knowledge" part of the model).

For developing high levels of EI skills, however, it is not sufficient to be motivated to enhance EI; it is also important to possess knowledge about the core facets of EI, such as knowledge of emotions and strategies for emotion regulation (see "Student" – "Motivation and Knowledge" in the model; cf., EI skills as outlined in the Investment Model; Zeidner, Matthews, Roberts, & MacCann, 2003). Knowledge of emotions *and* methods for regulating emotions are important components of instructional intervention with respect to enhancing learners' EI skills (see "Instructional Content" part of the model). Both aspects interact in their effect on students' knowledge of how to regulate emotions in a goal-directed way, similar to the interaction between controllability and value of emotions described previously.

Strategies for regulating emotions are only helpful if students have a clear vision of the learning or achievement goals that should be reached. Similarly, possessing knowledge of emotions is only helpful if students understand how to implement emotion-regulation strategies that use emotion knowledge to initiate goal-directed behavior. In other words, both aspects must reach a minimum level for students' knowledge of goal-directed emotion regulation to arise.

Learners' increased motivation to develop EI capacities, combined with greater knowledge of how to regulate emotions in a goal-directed way, should lead to enhanced EI skills (see the "Student" – "Skill" part of the model). According to the definition of EI implicated in the PEILAS model, EI skills reflect the ability to perceive, reflect upon, and regulate emotions. An interaction effect between motivation to enhance EI and knowledge of the goal-directed regulation of emotions can then be theoretically posited: high motivation to increase EI is only helpful if a minimum level of knowledge of the facets of EI exists and vice versa ("Student" – "Motivation and Knowledge" part of the model).

Finally, the PEILAS model views students' EI skills as recursively affecting their motivation to enhance EI (due to encouragement received in the learning environment) and their knowledge of how to regulate emotions in a goal-directed fashion (due to real-life learning experiences). Similarly, the instructional content pertaining to EI presented in the classroom will depend on students' baseline level of EI skills, which will continuously increase throughout instructional intervention aimed at fostering EI skills.

In the following section, we discuss strategies for enhancing students' EI based on the PEILAS model. We refer to, and expand upon, suggestions outlined in the description of the previous version of the model (see Goetz et al., 2005).

11.6.3 Fostering Emotional Intelligence According to the PEILAS Model

According to the PEILAS model, it is of chief importance *to teach learners that academic emotions are both controllable and valuable*. If academic emotions are viewed as uncontrollable, students' attempts at regulating their academic emotions will not be successful and consequently not meaningful – even if emotions are perceived as being highly important. Further, if academic emotions are not perceived as valuable, there will be less (or no) incentive to regulate them – even if academic emotions are seen as controllable. Thus, effective interventions must focus on both aspects. With respect to *teaching learners that emotions are controllable*, an educator or researcher might begin by asking students about the types of emotions they have experienced at school or other learning situations, and how stable they perceive these emotions to be (see corresponding literature on attribution theory, learned helplessness, and learned optimism; Möller & Köller, 1996; Peterson, 2000; Seligman, 1993, 2006; Weiner, 1985, 1995). A goal of this exercise is to help

students recognize the instability of emotional experiences and consequently that their academic emotions might be under their control. A specific technique for helping students reach this goal is emotion-focused retraining, an adopted version of traditional attributional retraining, which has typically focused on motivation rather than emotion (e.g., Perry, 1991; Schunk, 1984; Struthers, Perry, & Menec, 2000; van Overwalle & De Metsemaere, 1990). For example, if a student reports having experienced high levels of anxiety during their last math exam, one might respond by referencing the fact that the student may not have experienced equally high levels of anxiety when taking other math exams or when taking exams in other subject domains. This shows the student that his or her anxiety can be perceived (in part) as controllable and is not necessarily a stable trait (cf., Hall et al., 2007). There are also numerous ways to *teach learners that emotions are valuable*. Discussing with students the important effects academic emotions have on learning and achievement (see above) may increase – as a side effect – students' judgment of how valuable academic emotions are. Further, emphasizing the influence academic emotions have with respect to intrinsic motivation, subjective well-being, physical health (e.g., Ekman & Davidson, 1994; Goleman, 1995; Seligman, 2012; Zeidner et al., 2012) and job performance (see O'Boyle Jr., Humphrey, Pollack, Hawver, & Story, 2010) is expected to increase students' perceived importance of academic emotions.

As outlined in the PEILAS model, *teaching learners about emotions* and *teaching strategies for regulating emotions* is key to enhancing students' EI skills. Based on the definition of EI in this model, emotion knowledge includes knowledge about the nature of academic emotions and how they are classified, an emotion vocabulary, and most importantly, knowledge regarding how emotions affect learning processes and achievement. Strategies for regulating emotions include concrete ways, in which students regulate given emotions (i.e., response-focused emotion regulation, Gross, 1998) and are tied to students' emotion knowledge. In the following section, we suggest ways to foster students' knowledge of emotions and strategies for regulating emotions.

To increase students' knowledge of academic emotions, findings from emotion research can be used in the classroom to first *define academic emotions*. These findings can be modified with respect to students' age and previous knowledge of emotions. For example, definitions of academic emotions presented in this chapter can be used. Other researchers in the field present definitions of academic emotions that have varying degrees of complexity. For example, Frenzel and Stephens (2013) differentiate habitual (i.e., trait) and state emotions and distinguish emotions from related constructs like mood or stress. Frenzel and Stephens' (2013) chapter is intended primarily for teachers and teacher-students, and their (adopted) graphical depiction of the components of academic emotions could be used to facilitate classroom discussion – for example, selected emotions can be discussed in terms of their components.

Understanding and talking about emotions in a nuanced way require an emotion vocabulary that enables higher-level communication (Garner, 2010). There are numerous ways to *build students' emotion vocabulary*. One example is to discuss or

provide complex descriptions of the experience of "joy" by using adjectives like "cheery," "glad," or "bright." Extracurricular programs that focus on teaching emotional words also exist and can serve as a resource. An example is Greenberg and colleagues' (1995) PATHS program (see also Kusche & Greenberg, 2001), in which students learn emotion words using a hierarchical approach. That is, they begin with common emotions (e.g., happiness, sadness, anger) before moving on to learn about more complex emotional experiences (e.g., jealousy, guilt, pride). To consolidate the newly learned emotion vocabulary, students can be encouraged to reflect on the emotion words they use in daily life and in their writing.

It may also be helpful for students to learn how to *classify academic emotions*, in order to increase their skill in talking about and understanding their own emotions and the emotions of others. Emotion classification systems might help, for example, with discussing groups of emotions that share similarities based on specific features. Numerous approaches to classifying emotions exist, which can be adapted for use in the classroom by teachers. For example, Frenzel and Stephens (2013) present a system that classifies academic emotions based on the approach suggested by Pekrun and colleagues (e.g., Pekrun, Elliot, & Maier, 2006; Pekrun & Jerusalem, 1996). Among other criteria, academic emotions are grouped into positive (pleasant) emotions, such as enjoyment, hope, gratitude, or "Schadenfreude," and into negative (unpleasant) emotions, such as boredom, frustration, anger, fear, hopelessness, sadness, disappointment, shame, or guilt. This approach also differentiates between academic emotions according to object focus, specifically emotions that are directed primarily toward an activity (e.g., enjoyment of learning) vs. toward an outcome (e.g., enjoyment about good grades). One possibility for teaching emotion classification systems is to ask students to categorize various emotions and to outline the criteria for their arrangement. In the next step, categories suggested by students can be compared and contrasted with existing categories from the literature, such as those described above.

Students' degree of *knowledge about the effects of academic emotions* on learning and achievement is very important. Knowing which academic emotions are helpful and which are detrimental to learning is a prerequisite for goal-oriented emotion regulation. Academic emotion knowledge includes aspects such as how positive and negative emotions influence the way we learn and solve problems, or how emotions affect the activation of cognitive resources, motivation to learn, and achievement outcomes. For example, an adapted version of the model developed by Frenzel and Stephens (2013; based on Pekrun, 2006) illustrating the effects of academic emotions can be used for discussion and reflection with students. After a general discussion, students might focus on a specific emotion and how it affects learning and achievement. For example, academic boredom might be a particularly interesting emotion to discuss with students (for a detailed description of this emotion, see Goetz and Hall, 2014; for empirical results see Pekrun, Hall, Perry, & Goetz, 2014).

Teaching strategies for regulating academic emotions is a main facet of the PEILAS model. A starting point might be to discuss with students the meaning of "regulating" one's own emotions by referring to the term "self-regulation." This term can be introduced adapting theoretical models developed in the field of self-

regulated learning (for an overview see Goetz, Nett, & Hall, 2013). Students can be taught that the goal of self-regulation is to reduce the difference between an actual state and a target state in terms of individual learning or performance expectations. Students should understand that, from this perspective, self-regulation is strongly related to problem-solving (Anderson, 2000). In a next step, self-regulation might be discussed in the context of academic emotions. Students should understand that emotion-related self-regulation aims to change actual emotional states into target emotional states and that there are numerous ways to actively regulate one's own emotions to reach a specific goal. Depending on students' age and previous knowledge, specific strategies can be introduced and discussed. Examples of methods of emotion regulation include relaxation techniques (e.g., breathing techniques, autogenic training, and/or progressive muscle relaxation), positive self-instruction (self-communication, focusing on the controllability of academic emotions), and reducing the subjective experience of work-play dichotomy (Covington & Wiedenhaupt, 1997). However, introducing particular strategies for regulating academic emotions may be above and beyond the main competence or expectation of teachers. Thus, outside intervention professionals might be included in the intervention process to foster students' EI and academic emotions.

11.7 Conclusion

In this chapter we synthesized current research on academic emotions and presented definitions, approaches toward assessment, as well as antecedents and effects of this broad construct. Further, we attempted to motivate researchers, teachers, parents, and students to pay attention to and emphasize the importance of emotion regulation, especially in the academic context. Due to the fact that high EI can lead to a successful and goal-directed regulation of academic emotions, we introduced the PEILAS model. This model shows efficient ways of promoting emotional intelligence in learning and achievement situations.

References

Allen, V., MacCann, C., Matthews, G., & Roberts, R. D. (2014). Emotional intelligence in education: From pop to emerging science. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 162–182). New York: Routledge.

Amelang, M., Bartussek, D., Stemmler, G. & Hagemann, D. (2006). *Differentielle Psychologie und Persönlichkeitsforschung*. [Differential psychology and personality research] (6 ed.). Stuttgart: Kohlhammer.

Anderson, J. R. (2000). *Cognitive psychology and its implications*. New York: Freeman.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64*, 359–372.

Atkinson, J. W. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.

Barrett, L. F. (1998). Discrete emotions or dimensions? The role of valence focus and arousal focus. *Cognition and Emotion, 12*, 579–599. doi:10.1080/026999398379574.

Bieg, M., Goetz, T., & Hubbard, K. (2013). Can I master it and does it matter? An intraindividual analysis on control-value antecedents of trait and state academic emotions. *Learning and Individual Differences, 28*, 102–108.

Brackett, M. A., Lopes, P. N., Ivcevic, Z., Mayer, J. D., & Salovey, P. (2004). Integrating emotion and cognition: The role of emotional intelligence. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition. Integrative perspectives on intellectual functioning and development* (pp. 175–194). Mahwah, NJ: Lawrence Erlbaum Associates.

Brackett, M. A., & Rivers, S. E. (2014). Transforming students' lives with social and emotional learning. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 368–388). New York: Routledge.

Brunner, M., Keller, U., Dierendonck, C., Reichert, M., Ugen, S., Fischbach, A., et al. (2010). The structure of academic self-concepts revisited: The nested Marsh/Shavelson model. *Journal of Educational Psychology, 102*, 964–981. doi:10.1037/a0019644.

Cohen, J. (Ed.). (1999). *Educating minds and hearts: Social emotional learning and the passage into adolescence*. New York: Teachers College Press.

Cohen, J. (Ed.). (2001). *Caring classrooms/intelligent schools: The social emotional education of young children*. New York: Teachers College Press.

Cohn, J. F., Ambadar, Z., & Ekman, P. (2007). Observer-based measurement of facial expression with the facial action coding system. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 203–221). New York: Oxford University Press.

Covington, M. V., & Wiedenhaupt, S. (1997). Turning work into play: The nature and nurturing of intrinsic task engagement. In R. Perry & J. C. Smart (Eds.), *Effective teaching in higher education: Research and practice* (pp. 101–114). New York: Agathon Press.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *Journal of Nervous and Mental Disease, 175*, 526–536. doi:10.1097/00005053-198709000-00004.

D'Arcey, T., Johnson, M., & Ennis, M. (2012). Assessing the validity of FaceReader using facial electromyography. *Proceedings of APS 24th annual meeting.* Retrieved from www.darcey.us/pdf/facereader.pdf

Dodge, K. A. (1989). Coordinating responses to aversive stimuli: Introduction to a special section on the development of emotion regulation. *Developmental Psychology, 25*, 339–342. doi:10.1037/0012-1649.25.3.339.

Dweck, C. S. (2002). Messages that motivate: How praise molds students' beliefs, motivation, and performance (in surprising ways). In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 37–60). San Diego: Academic.

Eid, M., Schneider, C., & Schwenkmezger, P. (1999). Do you feel better or worse? The validity of perceived deviations of mood states from mood traits. *European Journal of Personality, 13*(4), 283–306.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial action coding system [e-book]*. Salt Lake City, UT: Research Nexus.

Ekman, P., & Davidson, R. J. (1994). Affective science: A research agenda. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 411–434). New York: Oxford University Press.

Ekman, P., & Rosenberg, E. L. (2005). *What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system* (FACS, 2nd ed.). New York: Oxford University Press.

Elias, M. J., Hunter, L., & Kress, J. S. (2001). Emotional intelligence and education. In J. Ciarrochi, J. P. Forgas, & J. D. Mayer (Eds.), *Emotional intelligence in everyday life* (pp. 133–149). Philadelphia: Psychology Press.

Frenzel, A. C., & Stephens, E. J. (2013). Emotions. In N. C. Hall & T. Goetz (Eds.), *Emotion, motivation, and self-regulation: A handbook for teachers* (pp. 1–56). Bingley, UK: Emerald.

Galassi, J. P., Frierson, H. T., & Sharer, R. (1981). Behavior of high, moderate, and low test anxious students during an actual test situation. *Journal of Consulting and Clinical Psychology, 49*, 51–62. doi:10.1037/0022-006x.49.1.51.

Garner, P. W. (2010). Emotional competence and its influences on teaching and learning. *Educational Psychological Review, 22*, 297–321.

Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R., & Hall, N. C. (2013). Do girls really experience more anxiety in mathematics? *Psychological Science, 24*, 2079–2087. doi:10.1177/0956797613486989.

Goetz, T., Frenzel, A. C., Hall, N. C., Nett, U., Pekrun, R., & Lipnevich, A. (2014). Types of boredom: An experience sampling approach. *Motivation and Emotion, 38*, 401–419.

Goetz, T., Frenzel, C. A., Pekrun, R., & Hall, N. (2005). Emotional intelligence in the context of learning and achievement. In R. Schulze & R. D. Roberts (Eds.), *Emotional intelligence: An international handbook* (pp. 233–253). Cambridge, MA: Hogrefe & Huber Publishers.

Goetz, T., Frenzel, A. C., Stoeger, H., & Hall, N. C. (2010). Antecedents of everyday positive emotions: An experience sampling analysis. *Motivation and Emotion, 34*, 49–62.

Goetz, T., & Hall, N. C. (2013). Emotion and achievement in the classroom. In J. Hattie & E. M. Anderman (Eds.), *International guide to student achievement* (pp. 192–195). New York: Taylor & Francis.

Goetz, T., & Hall, N. C. (2014). Academic boredom. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 311–330). New York: Routledge.

Goetz, T., Nett, U. E., & Hall, N. C. (2013). Self-regulated learning. In N. C. Hall & T. Goetz (Eds.), *Emotion, motivation, and self-regulation: A handbook for teachers* (pp. 123–166). Bingley, UK: Emerald.

Goetz, T., Pekrun, R., Hall, N. C., & Haag, L. (2006). Academic emotions from a social-cognitive perspective: Antecedents and domain specificity of students' affect in the context of Latin instruction. *British Journal of Educational Psychology, 76*, 289–308.

Goleman, D. (1995). *Emotional intelligence: Why it can matter more than IQ*. New York: Bantam Books.

Graham, S., & Taylor, A. Z. (2014). An attributional approach to emotional life in the classroom. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 96–119). New York: Routledge.

Greenberg, M. T., Kusche, C. A., Cook, E. T., & Quamma, J. P. (1995). Promoting emotional competence in school-aged children: The effects of the PATHS curriculum. *Development and Psychopathology, 7*, 117–136.

Gross, J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*, 271–299.

Gross, J. J. (Ed.). (2007). *Handbook of emotion regulation*. New York: The Guilford Press.

Hall, N. C., Perry, R. P., Goetz, T., Ruthig, J. C., Stupnisky, R. H., & Newall, N. E. (2007). Attributional retraining and elaborative learning: Improving academic development through writing-based interventions. *Learning and Individual Differences, 17*, 280–290.

Humphrey, N., Curran, A., Morris, E., Farrell, P., & Woods, K. (2007). Emotional intelligence and education: A critical review. *Educational Psychology, 27*, 235–254.

Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: An integrative meta-analysis and cascading model. *Journal of Applied Psychology, 95*, 54–78.

Kusche, C. A., & Greenberg, M. T. (2001). PATHS in your classroom: Promoting emotional literacy and alleviating emotional distress. In J. Cohen (Ed.), *Caring classrooms/intelligent schools: The social emotional education of young children* (pp. 140–161). New York: Teachers College Press.

Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), *Technology in mental health care delivery systems* (pp. 119–137). Norwood, NJ: Ablex.

Lau, P. S. Y., & Wu, F. K. Y. (2012). Emotional competence as a positive youth development construct: A conceptual review. *The Scientific World Journal*, *2012*, 8 pages, Article ID 975189.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.

MacCann, C., Joseph, D. L., Newman, D. A., & Roberts, R. D. (2014). Emotional intelligence is a second-stratum factor of intelligence: Evidence from hierarchical and bifactor models. *Emotion, 14*(2), 358–374.

Matthews, G., Roberts, R. D., & Zeidner, M. (2004). Seven myths about emotional intelligence. *Psychological Inquiry, 15*, 179–196.

Matthews, G., Zeidner, M., & Roberts, R. D. (2002). *Emotional intelligence: Science and myth*. Cambridge, MA: MIT Press.

Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology, 59*, 507–536.

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. Sluyter (Eds.), *Emotional development and emotional intelligence: Implications for educators* (pp. 3–31). New York: Basic Books.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2002). *Mayer-Salovey-Caruso emotional intelligence test MSCEIT user's manual*. North Tonawanda, NY: Multi-Health Systems.

Meinhardt, J., & Pekrun, R. (2003). Attentional resource allocation to emotional events: An ERP study. *Cognition and Emotion, 17*, 477.

Möller, J., & Köller, O. (1996). Attributionen und Schulleistung [Attribution and scholastic performance]. In J. Möller & O. Köller (Eds.), *Emotionen, Kognitionen und Schulleistung [Emotions, cognitions and scholastic achievement]* (pp. 115–136). Weinheim: Psychologie Verlags Union.

Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K.-T., & Trautwein, U. (2011). Who took the "×" out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science, 22*, 1058–1066. doi:10.1177/0956797611415540.

Noldus Information Technology. (2012). *FaceReader™ reference manual version 5*. Retrieved from http://www.noldus.com/webfm_send/1230

O'Boyle, E. H., Jr., Humphrey, R. H., Pollack, J. M., Hawver, T. H., & Story, P. (2010). The relation between emotional intelligence and job performance: A meta-analysis. *Journal of Organizational Behavior, 35*, 788–818.

Parker, J. D. A., Saklofske, D. H., Wood, L. M., & Collin, T. (2009). The role of emotional intelligence in education. In C. Stough, D. H. Saklofske, & J. D. A. Parker (Eds.), *Assessing emotional intelligence. Theory, research and applications* (pp. 239–255). New York: Springer.

Pekrun, R. (1992). The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Applied Psychology: An International Review, 41*(4), 359–376. doi:10.1111/j.1464-0597.1992.tb00712.x.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315–341. doi:10.1007/s10648-006-9029-9.

Pekrun, R., & Bühner, M. (2014). Self-report measures of academic emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 561–579). New York: Routledge.

Pekrun, R., Elliot, A. J., & Maier, M. A. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *Journal of Educational Psychology, 98*, 583–597. doi:10.1037/0022-0663.98.3.583.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*, 36–48. doi:10.1016/j.cedpsych.2010.10.002.

Pekrun, R., Hall, N. C., Goetz, T., & Perry, R. P. (2014). Boredom and academic achievement: Testing a model of reciprocal causation. *Journal of Educational Psychology, 106*(3), 696–710.

Pekrun, R., & Jerusalem, M. (1996). Leistungsbezogenes Denken und Fühlen: Eine Übersicht zur psychologischen Forschung [Cognitions and emotions related to achievement: A review on psychological research]. In J. Möller & O. Köller (Eds.), *Emotionen, Kognitionen und Schulleistung [Emotion, cognition, and scholastic achievement]* (pp. 3–22). Weinheim: Psychologie Verlags Union.

Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 120–141). New York: Routledge.

Perry, R. P. (1991). Perceived control in college students: Implications for instruction in higher education. In J. Smart (Ed.), *Higher education: Handbook for theory and research* (Vol. 7, pp. 1–56). New York: Agathon Press.

Peterson, C. (2000). The future of optimism. *American Psychologist, 55*(1), 44–55.

Reisenzein, R., Junge, M., Studtmann, M., & Huber, O. (2014). Observational approaches to the measurement of emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 580–606). New York: Routledge.

Rivers, S. E., Brackett, M. A., Reyes, M. R., Mayer, J. D., Caruso, D. R., & Salovey, P. (2012). Measuring emotional intelligence in early adolescence with the MSCEIT-YV: Psychometric properties and relationship with academic performance and psychosocial functioning. *Journal of Psychoeducational Assessment, 30*, 344–366.

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128*, 934–960. doi:10.1037/0033-2909.128.6.934.

Rost, D. H., & Schermer, F. J. (1987). Emotion and cognition in coping with test anxiety. *Communication and Cognition, 20*, 225–244.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178. doi:10.1037/h0077714.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality, 9*, 185–211.

Schober, B., Finsterwald, M., Wagner, P., Lüftenegger, M., Aysner, M., & Spiel, C. (2007). TALK – A training program to encourage lifelong learning in school. *Journal of Psychology, 215*, 183–193.

Schunk, D. H. (1984). Sequential attributional feedback and children's achievement behaviors. *Journal of Educational Psychology, 76*, 1159–1169.

Seligman, M. E. P. (1993). *Learned helplessness*. New York: Oxford University Press.

Seligman, M. E. P. (2006). *Learned optimism. How to change your mind and your life*. New York: Vintage Books.

Seligman, M. E. P. (2012). *Flourish. A visionary new understanding of happiness and well-being*. New York: Free Press.

Spielberger, C., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the state-trait anxiety inventory*. Palo Alto, CA: Consulting Psychologists Press.

Struthers, C. W., Perry, R. P., & Menec, V. H. (2000). An examination of the relationships among academic stress, coping, motivation, and performance at college. *Research in Higher Education, 41*, 579–590.

Topping, K. J., Holmes, E. A., & Bremner, W. G. (2000). The effectiveness of school-based programs: For the promotion of social competence. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence* (pp. 411–432). San Francisco: Jossey-Bass.

van Overwalle, F., & De Metsenaere, M. (1990). The effects of attribution-based intervention and study strategy training on academic achievement in college freshmen. *British Journal of Educational Psychology, 60*, 299–311.

Vansteelandt, K., van Mechelen, I., & Nezlek, J. B. (2005). The co-occurrence of emotions in daily life: A multilevel approach. *Journal of Research in Personality, 39*, 325–335. doi:10.1016/j.jrp.2004.05.006.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070. doi:10.1037/0022-3514.54.6.1063.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*, 548–573. doi:10.1037/0033-295x.92.4.548.

Weiner, B. (1995). *Judgements of responsibility: A foundation for a theory of social conduct*. New York: Guildford Press.

Zeidner, M. (1998). *Test anxiety: The state of the art*. New York: Plenum Press.

Zeidner, M., & Endler, N. (Eds.). (1996). *Handbook of coping*. New York: Wiley.

Zeidner, M., Matthews, G., & Roberts, R. D. (2012). The emotional intelligence, health, and well-being nexus: What have we learned and what have we missed? *Applied Psychology Health and Well-being, 4*, 1–30.

Zeidner, M., Matthews, G., Roberts, R. D., & MacCann, C. (2003). Development of emotional intelligence: Towards a multi-level investment model. *Human Development, 46*, 69–96. doi:10.1159/000068580.

Part III
Psychosocial Skills: Applications

Chapter 12
Fostering Psychosocial Skills: School-Based Promotion of Resiliency in Children and Adolescents

Sandra Prince-Embury, Kateryna V. Keefer, and Donald H. Saklofske

12.1 School-Based Promotion of Resiliency in Children and Adolescents

Resilience in the face of adversity has been studied extensively by psychologists for the past 50 years. This body of work has defined the common theme of resilience as the ability to weather adversity or to bounce back from a negative experience. Research on resilience suggests that psychological symptoms and disorders may be based in part on lower personal resiliency or greater vulnerability to situations and events that the person has experienced (Garmezy, 1971, 1985, 1991; Garmezy, Masten, & Tellegen, 1984; Luthar, 1991; Luthar, Cicchetti, & Becker, 2000; Luthar & Zigler, 1991, 1992; Masten, 2001; Masten & Coatsworth, 1998; Masten & Curtis, 2000; Masten & Powell, 2003; Masten et al., 2005; Prince-Embury, 2007, 2008, 2013a; Prince-Embury & Saklofske, 2013, 2014; Rutter, 1987, 1993).

The definition of resilience as a product of complex interactions of personal attributes and environmental circumstances, mediated by internal mechanisms, has presented a challenge to those interested in applying the construct to human behavior in everyday and extreme circumstances (Luthar et al., 2000). In an effort to clarify constructs, theorists have distinguished "resilience" from "resiliency"; the former is defined as interactive and contextual and the latter addresses personal attributes of the

S. Prince-Embury (✉)
The Resiliency Institute of Allenhurst, LLC., 625 North Edgemere Drive, West Allenhurst, NJ, USA
e-mail: Sandraprince-embury@earthlink.net

K.V. Keefer • D.H. Saklofske
Department of Psychology, University of Western Ontario, Social Science Centre, 7th Floor, London, ON, Canada
e-mail: kvkeefer@gmail.com; dsaklofs@uwo.ca

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_12

individual (Luthar et al., 2000; Luthar & Zelazo, 2003; Masten & Coatsworth, 1998). Some resilience research has employed longitudinal studies, reflecting a developmental perspective, and tried to capture contextual aspects of resilience specific to groups and sets of circumstances. Studies assessing personal resiliency, in an effort to be comprehensive, have employed extensive assessment batteries, along with various criteria of competence, achievement, or successful adaptation (Werner & Smith, 1982).

Earlier research findings on resilience were interpreted to suggest that resilient individuals are "extraordinary" and that this quality or characteristic is not accessible to everyone. More recently, Masten described resilience as the process characteristic of normal development, an "ordinary magic," and not just applicable in adverse circumstances (Masten, 2001; Masten & Powell, 2003). Masten (2001) suggested that fundamental systems, already identified as characteristic of human functioning, have great adaptive significance across diverse stressors and threatening situations. This shift in emphasis had significant implications, and the "ordinary magic" framework suggested by Masten led to the extended application of resilience theory to a wider range of individuals in varied contexts. These systems include attachment relationships and social support; intelligence and problem-solving skills; self-regulation skills involved in directing or inhibiting attention, emotion, and action (Chap. 9); agency, mastery, motivation, and self-efficacy (Chap. 10); *meaning making* (constructing meaning and a sense of coherence in life); and cultural traditions.

12.1.1 The Role of Resilience in Schools

In spite of its conceptual complexity, "resilience" has been recognized by educators and school psychologists as a concept that is consistent with overall educational goals and well suited for application in educational settings. Resilience has been applied in more education-specific ways as "academic resilience." Academic resilience may be defined as the ability to effectively deal with setbacks, stress, or pressure in an academic or school setting. This concept has been employed to understand academic success in poor, minority, and disadvantaged students (Wang, 1994). Rutter (1987), in early discussions of resilience, identified four types of mechanisms that are applicable in academic settings for mediation of adverse circumstances: reducing the impacts of risks, reducing the likelihood of negative chain reactions associated with adversity, establishing and maintaining self-esteem and self-efficacy, and creating new opportunities for success. Examples of the applications of Rutter's four mechanisms are already in place in many academic settings. Children from impoverished families may attend "Head Start" or early enrichment programs to better prepare them for success in regular school classrooms. Increasingly, educators have focused on identifying learning needs and employed a "strength-based" perspective rather than a "disability" perspective which leaves the child to feel inadequate and disempowered. An attitude of respect and both personal and academic self-worth may be fostered by individual teachers and school

environments in general. Learning opportunities are offered in various formats to accommodate the diversity of strengths and learning styles in children, drawing from what we know about the characteristics of learners and fitting this to the most effective learning environments.

The National Association of School Psychologists embraced the theme of "Resilience: Building Strength for Life" for its 2008 Conference in New Orleans. This theme was part of a year-long initiative aimed at integrating the resilience concept into the practice of school psychology and presenting practices to build resilience within the school setting. Presentations suggested that application of resilience/resiliency constructs in educational environments made sense for many reasons. The constructs are based on relative strength and vulnerability as opposed to a deficit model or clinical pathology. The constructs relate to academic achievement and positive educational environments as well as avoidance of pathology and dysfunction. The constructs are developmental and normative and may be applied universally to guide system-level practice as well as individually to screen for children and youth who may be at risk. School psychologist and other practitioners have already begun applying principles of resilience to education (Doll, Zucker, & Brehm, 2004; see also Prince-Embury & Saklofske, 2014) in the classroom (Brooks & Goldstein, 2001), by coaching parents (Brooks & Goldstein, 2001) and by coaching teachers (Brooks & Goldstein, 2008). Resilience in education has also been linked with school-based mental health initiatives intended to create environments and the developmental skills that encourage the psychological well-being of children and that further extend into the community at large (see Canadian Journal of School Psychology, 2013, 28, 1).

12.2 Conceptualizing Resiliency for Applications in School Settings

Many concepts and many interventions are subsumed under the umbrella constructs of resilience/resiliency, including constructs of social-emotional learning (SEL). This has led some to argue that lack of consensus on a definition limits any practical use of resilience in understanding, predicting, and changing human behavior (Kaplan, 1999, 2005). Alternatively, some have claimed that in spite of conceptual complexity, the phenomenon of resilience has too much heuristic power to be abandoned (Luthar et al., 2000). Elias, Parker, and Rosenblatt (2005) proposed the use of working definitions of resilience/resiliency that satisfy two criteria: (1) does the definition add value to existing constructs in understanding circumstances; (2) does the definition inform the design of interventions. Kaplan's (2005) review conceded that concepts are not by their nature true or false but may be evaluated with regard to their usefulness. Given such conceptual debates, one might ask what added value is offered by the construct of resiliency over the construct of SEL, or conversely are these really the same? We suggest that although there is much overlap between resiliency and SEL, there are differences between the two. Social-emotional learning

enhances the relational and self-regulatory skills of children as assets that enhance their ability to function and learn in an educational setting as well as future settings. Resiliency is comprised of core factors that help children in the face of possible adversities, large and small, that might be encountered during their school years and beyond. Thus, while social-emotional competency can be viewed as a facet of resiliency, in that resiliency includes relational and self-regulatory skills, resiliency also includes other factors that are important in the presence of adversity.

12.2.1 Three-Factor Model of Personal Resiliency

One effort to simplify the construct of personal resiliency for non-stigmatizing assessment and application in the schools is the three-factor model developed by Prince-Embury (2007, 2013a, 2013b, 2014). This model is based on three previously identified attributes of personal resiliency reflective of three core developmental systems, sense of mastery, sense of relatedness, and emotional reactivity, and the relationship of these factors to one another (Prince-Embury, 2006, 2007, 2013a, 2013b, 2014). Earlier models of personal resiliency suggested one factor: resiliency as a trait was viewed by some as present to some degree or not (Block, 1980, 2002). Other resiliency literatures discussed two factors: protective or risk. The three-factor model was based on review of the literature, clinical practice, and factor analysis (Prince-Embury, 2007, 2013a, 2013b; Prince-Embury & Courville, 2008a, 2008b). The three-factor model includes many of the constructs discussed in association with SEL but groups them into three conceptual categories for the purpose of clearer needs assessment, intervention, and outcomes assessment. Unlike the skill-based SEL model, the three-factor model of personal resiliency focuses on the personal experience of the child and not actual ability or performance as assessed by others. Although it is recognized that actual ability and performance as assessed by others is important, the three-factor model assumes that the child's experience mediates between external protective factors and positive behavioral outcomes. The definition of resiliency according to underlying developmental systems is designed to aid in the identification of appropriate interventions that may be most needed at the individual or aggregate level. Definition and application of each of the three factors of personal resiliency are described briefly below.

Sense of Mastery One set of core mechanisms that have been consistently identified as important for resiliency in developmental and resilience research are sense of mastery and self-efficacy. White (1959) suggested that children's sense of competence or efficacy provides them with the opportunity to interact with and enjoy cause-and-effect relationships in the environment. Bandura (1977, 1993) suggested that students' self-efficacy beliefs for regulating their own learning and mastering academic activities determine their aspirations, level of motivation, and academic accomplishments (see also Chap. 9). Positive expectations about their future predicted lower anxiety, higher school achievement, and better classroom behavior

control (Wyman, Cowen, Work, & Kerley, 1993). Previous research and theory suggests that children and youth who have a greater sense of competence/efficacy may be more likely to succeed in a school environment and less likely to develop pathological symptoms. Interventions to enhance sense of mastery have significant implications for the school setting. Enhanced and realistic sense of mastery increases students' expectations and attempts to achieve these expectations, which in turn may enhance a sense of mastery. Looking at the school environment contextually, we refer to several specific pathways for enhancing sense of mastery: lessons that are matched to the ability level of students and broken into achievable steps, reducing the likelihood of negative chain reactions associated with adversity, establishing and maintaining self-esteem and self-efficacy, and creating new opportunities for success (Rutter, 1987, 1993, 2010); teachers trained to foster a resilient mindset in students (Brooks & Goldstein, 2008); a teaching process that redefines "failure" as overcoming challenges and problem solving; and classroom environments that are responsive to the feedback of students in creating a more resilient classroom (Doll et al., 2004).

Sense of Relatedness Reviewing five decades of resilience research in child development, Luthar (2006, p. 780) concluded, "Resilience rests, fundamentally, on relationships." The importance of relationships for human resilience has been noted in every major review of protective factors for resilience (see Masten & Obradovic 2006). The importance of relationships and relational ability as mediators of resilience has been supported in research by developmental psychologists. Much developmental theory has been devoted to the development of internal mechanisms of attachment and relatedness (see Prince-Embury, 2007 for discussion). Werner and Smith (1982) noted that resilient youth sought support from nonparental adults (especially teachers, ministers, and neighbors) more often than non-resilient youth. It must be noted, however, that previous research has indicated that perceived support, as distinguished from actual support, is the dimension of social support that is most strongly related to psychological well-being in adults and children (see Prince-Embury, 2007 for discussion). Efforts to enhance actual relatedness and perceived support have focused on enhancement of social skills through social-emotional learning. The logic is that youth with better social skill will have better relationships and enhanced sense of relatedness.

Within the context of social-emotional learning, much thought and effort has been given to enhancing social skills in children such as communication, cooperation, assertion, empathy, engagement, and self-control, which may be broken down into teachable skills such as improving eye contact, initiating and maintaining conversations, understanding others' feelings and promoting empathy, sharing, and maintaining personal space (Alvord, Zucker, & Grados, 2011; de Boo & Prins, 2007). Such programs as Social Skills Improvement System (Gresham & Elliot, 1990) have been successfully used with children and adolescents to increase interpersonal competencies when these are lacking and are among the major contributors to a child's social and emotional difficulties.

Enhancing interpersonal skills in youth may enhance school engagement and performance and perhaps more general sense of relatedness in the long run. The implication for application in the school environment is that better social skills increase the likelihood of better social relationships with peers and teachers and less conflict that interferes with learning and school attendance. In addition, research has indicated that better social engagement in school is associated with better academic performance. Students who have friends at school are more interested in academic activities and are more active participants in the classroom (Malecki & Elliott, 2002; Wentzel & Watkins, 2002). This is consistent with the assumption that learning and achievement takes place within a meaningful social context and that strength of engagement of students with teachers and other students indicates the social meaningfulness of the school environment. In summary, research suggests that a positive sense of relatedness within the school environment is essential for meaningful learning and academic achievement. Therefore, efforts to enhance students' social engagement with peers and teachers within the school environment would enhance the educational goals of the school. In addition, focus on sense of relatedness and other SEL skills is also consistent with trait models of emotional intelligence (EI), which suggest that social-emotional competencies have positive implications throughout the lifespan (Bar-On, 2006; Petrides, 2011).

Emotional Reactivity Developmental research has demonstrated that children's development of pathology in the presence of adversity is related to their emotional reactivity and their inability to regulate this reactivity (Prince-Embury, 2013b). Specifically, strong emotional reactivity and related difficulty with regulation of this reactivity have been associated with behavioral maladjustment and vulnerability to pathology. Emotional reactivity is in part the child's arousability or the threshold of tolerance that exists prior to the occurrence of adverse events or circumstances. Rothbart and Derryberry (1981) have defined emotional reactivity as the speed and intensity of a child's negative emotional response. Children's reactivity varies in its intensity, sensitivity, specificity, windows of tolerance, and recovery (Siegel, 1999). Conversely, emotional regulation, or the ability to modulate emotional responses, is a significant factor in fostering resilience (Cicchetti, Ganiban, & Barnett, 1991; Cicchetti & Tucker, 1994; Eisenberg, Champion, & Ma, 2004). Regulation and redirection of emotional arousal is necessary for children to function adaptively in emotionally challenging situations (Cicchetti et al., 1991; Thompson, 1990).

Emotional reactivity in the school environment may be viewed as a source of impaired functioning and thus an impediment to learning. Importantly, academic achievement and behavioral self-control are highly interdependent. Students who are attentive, regulated, and persistent in their work often earn higher grades, whereas those who lack behavioral self-control often underachieve academically (Doll et al., 2004). Some studies have found disciplined classroom behavior to be a better predictor of students' grades than intellectual ability (McDermott, Mordell, & Stoltzfus, 2001). Existing programs to address emotional reactivity in school environments may involve relaxation exercises; learning how to accurately identify, label, and verbalize emotions; and regular opportunities to discharge excess energy.

The significance of emotional reactivity and emotion regulation in the school environment may be viewed on many levels. First, individual differences in students' physiologically based emotional reactivity may make adaptation to a structured, sedentary school environment difficult for those with higher emotional reactivity. Interventions for such children may involve behavior management and relaxation techniques or in some cases medication prescribed by physicians to lower base emotional reactivity. On the level of the school environment itself, we may examine potential triggers of emotional reactivity for children in general. Triggers may include novelty such as starting a new school or transitioning from elementary to middle, or middle to high school; presentation of material at a level too difficult for the student; punitive consequences for difficulties in learning; and difficulties in peer relationships including but not limited to bullying. Interventions in these instances would involve identifying triggers of emotional reactivity, preparation for these triggers, and efforts to modify these triggers to more emotionally neutral events. In summary, enhancing school resilience through addressing emotional reactivity might involve the following: identifying youth with higher emotional reactivity, teaching students to recognize early signs of emotional reactivity, and teaching them techniques to self-regulate and manage emotions, reducing the potential of environmental triggers to increase emotional reactivity in the school environment.

12.2.2 Need for Resiliency Assessment in the Schools

Assessment is the cornerstone of effective intervention. Studies of resilience have been both cross-sectional and longitudinal, have employed a developmental-psychopathology perspective, and have tried to capture contextual aspects of resilience specific to the group and sets of circumstances. Researchers of both resilience and resiliency have used different measures across studies and across populations making it difficult to compare across studies and across groups. The resiliency measures employed in research have often been impractical for widespread use in the school community because they are too labor intensive or expensive. On the other hand, some measures are restricted in their definition of resiliency or may not be linked with current or identifiable models of resiliency. From a psychometric perspective, some measures have less than adequate reliability and validity and may not have gone through the kind of standardization that would provide normative data that aid in the interpretation of an individual's scores relative to peers or clinical groups. The lack of common metrics across different studies of resilience/resiliency constructs and across research and practice results in difficulty assessing the effectiveness of intervention strategies in a way that allows comparison across methods and populations (Prince-Embury, 2011).

On a practical level, there is work to be done to make resiliency assessment tools more field friendly (Masten, 2001; Masten & Powell, 2003). Hence, there is a need for measures and benchmarks describing resiliency that are brief, easily administered, and simple to score and interpret. In addition, measures used with

diverse school populations must be bias-free with respect to gender and ethnicity and worded so that they might be used with a broad range of reading levels. In order to be acceptable to parents, students, and teachers in school settings, a measure assessing resiliency needs to be strength based and informative while at the same time not stigmatizing or "pathologizing" of groups or individuals (Prince-Embury, 2011).

Resiliency Scales for Children and Adolescents Prince-Embury developed the *Resiliency Scales for Children and Adolescents* (RSCA; Prince-Embury, 2006, 2007) for use in preventive universal screening to identify areas of strength and vulnerability at the aggregate and individual level, for planning resiliency enhancing interventions in the schools (Prince-Embury, 2010). The RSCA consists of three global scales based on the three-factor model of personal resiliency discussed above: sense of mastery, sense of relatedness, and emotional reactivity. Each of the global scales is further composed of several subscales: sense of mastery includes optimism, self-efficacy, and adaptability; sense of relatedness encompasses trust, comfort with others, support, and tolerance; emotional reactivity comprises sensitivity, recovery, and impairment. The RSCA is completed by the child (self-report) and written at a third-grade reading level and takes 10 min to complete.

The three global scale scores (mastery, relatedness, and emotional reactivity) may be used to plot each child's Personal Resiliency Profile, which highlights the individual child's relative strengths (mastery and/or relatedness) and vulnerability (emotional reactivity). At an individual level, the Personal Resiliency Profile may be used to guide the selection of an intervention or treatment plan. For example, youth who are low in sense of mastery may be presented with gradual achievable tasks toward specific educational goals. Youth with low sense of relatedness may be offered social skill training. Youth with high emotional reactivity may be presented with relaxation exercises and self-regulation skill training.

Examination of individual and aggregate Personal Resiliency Profiles indicated that although there was considerable individual variability, the two protective factors, mastery and relatedness, were often correlated with each other and negatively correlated with emotional reactivity (Prince-Embury, 2007, 2013a, 2013b). For this reason it is possible to condense the three-factor scores into two index scores for screening. The two protective scores, mastery and relatedness, may be combined to form a resource index score (see Prince-Embury, 2007 for details). Vulnerability then may be represented as the discrepancy between the emotional reactivity score and the resource index score (see Prince-Embury, 2007, for details). These two RSCA index scores, resource and vulnerability, may then be used for preventive, non-pathologizing screening in school systems.

The RSCA was standardized for three age groups (9–11, 12–14, 15–18) and stratified by ethnicity and parent education level within age group and gender. The RSCA scores demonstrate good to excellent reliability at the index, global scale, and subscale levels. Also, convergent and divergent validity evidence has been demonstrated (Prince-Embury, 2006, 2007, 2008, 2010).

Multitiered Screening Using the RSCA Index Scores A preventive screening model using the RSCA index scores (Prince-Embury, 2010) was presented in Doll, Pfohl, and Yoon's *Handbook for Youth Prevention Science* (2010). This model is briefly described below:

1. First Tier: Administer the RSCA on a school-wide or class-wide level and calculate global scale and index *T*-scores (see Prince-Embury, 2010).
2. Second Tier: If the vulnerability index score is *T*60 (high) or higher and if the emotional reactivity score is *T*60 or higher, then students may be identified for preventive intervention addressing management of emotional reactivity.
3. Third Tier: If the resource index is *T*40 (low) or below, examine the sense of mastery and sense of relatedness scale scores to determine specific areas for preventive intervention.
   (a) If the sense of mastery score is *T*40 or below, refer for preventive intervention pertaining to sense of mastery, self-efficacy, and adaptability.
   (b) If the sense of relatedness is *T*40 or below, refer for preventive intervention pertaining to sense of relatedness, social skills, communication skills, etc.

This preventive screening model begins with the vulnerability index score and follows up with the emotional reactivity scale score and the resource index score, taking the steps indicated above. To illustrate how the RSCA Index and scale scores might be used for screening, the following example is provided. Estimated numbers of students are based on cumulative percentages of scores obtained in the normative sample. If the RSCA was administered to a school population of 1000 at the beginning of the academic year, Tier 2 might identify 130 students (13 %) as having vulnerability index scores equal to or greater than *T*60, identifying them as potentially high in vulnerability and warranting preventive intervention. Of this identified group, many might also have emotional reactivity *T*-scores equal to or greater than *T*60. This group would be identified as potentially high in emotional reactivity and potentially in need of preventive intervention aimed at reducing emotional reactivity. As indicated in Tier 3, 110 students (11 % of total) would have resource index scores equal to or less than *T*40, suggesting that these students are low in resources and warrant preventive intervention to enhance resources. Approximately 85 (8.5 % of total) students might meet both criteria: vulnerability index and emotional reactivity scores of *T*60 or above, as well as resource index scores of *T*40 or below. Preventive intervention services might be offered based on the availability of resources.

The chapter thus far has discussed a conceptual model for assessment and application organized by developmental principles underlying personal resiliency. The second half of the chapter presents specific assessment and intervention approaches considered at different levels of implementation: school or system, classroom, and individual.

12.3 Resiliency Interventions at Different Levels of Implementation

The basic assumption underlying resilience-based interventions with children and youth in school contexts is that resiliency is not an immutable trait or end outcome, but rather a competency, a cognitive-behavioral style that can be learned by students and cultivated through supportive school and classroom environments (Prince-Embury, 2013a). From a practical standpoint, the view of resilience as "the everyday magic of ordinary, normative human resources" (Masten, 2001, p. 235) is applicable to all students. Apart from informing interventions for children who are already experiencing elevated risk or adversity, the recognition that all youth, regardless of their current circumstances, can benefit from developing greater resiliency and capacity to manage and adapt to their world gives schools an opportunity to engage in preventative action long before risks accumulate or problems develop (Brooks & Brooks, 2014; Mallin, Walker, & Levin, 2013).

12.3.1 Applications at the School or Systems Level

Given their central role in children's education and socialization, schools are the ideal venue for large-scale preventative efforts to promote children's resiliency and adaptation (Mallin et al., 2013; Schwean & Rodger, 2013). Although schools routinely implement evidence-based character-building and mental health promotion programs, many of them have not been systematically linked to the core resiliency constructs or assessment models (Prince-Embury & Saklofske, 2013, 2014). In this regard, social and emotional learning (SEL) programs represent a notable exception. As discussed previously, SEL and personal resiliency are overlapping concepts. These constructs differ in that resiliency is more often viewed in the context of adversity. Thus, resiliency was probably seen as less applicable in school environments as some definitions required the presence of an adverse context or circumstances. However, with the increased understanding of resilience as "ordinary magic" has come an increased understanding of adversity as everyday experience as well as obvious tragedy. For example, transition to a new school, bullying, being gay, or loss of a loved one require resiliency for many students. It is likely that the SEL rubric was more compatible with application in the school environment in that it framed social and emotional constructs as learnable skills, much like academic subjects, and thus appropriate for a learning environment. Similarly, resiliency as the ability to overcome obstacles may be broken down into teachable steps within the school environment.

The SEL approach to education is premised on the recognition that students' academic outcomes (i.e., school engagement, mastery of material, academic achievement) are significantly tied to their emotional and interpersonal functioning, in that when students feel overwhelmed emotionally or maladjusted socially, their

capacity to participate in and benefit from academic schooling also suffers (Greenberg et al., 2003; Zins, Bloodworth, Weissberg, & Walberg, 2004). The goal of SEL programs is to improve students' emotional, social, and academic outcomes by strengthening their personal resiliency resources in the emotional and relatedness domains (Merrell & Gueldner, 2010). The core socioemotional competencies targeted by SEL programs include self-awareness, self-management, social awareness, relationship skills, and responsible decision making. Thus, students are taught specific skills required for understanding and regulating one's emotions and behaviors, identifying and capitalizing on personal strengths and weaknesses, setting and achieving personal and academic goals, feeling and expressing empathy for others, establishing and maintaining rewarding interpersonal relationships, and making socially conscious choices and decisions (Collaborative for Academic, Social, and Emotional Learning [CASEL], 2013).

A recent meta-analysis of 213 controlled studies evaluating the outcomes of universal school-based SEL curricula found that well-implemented SEL programs resulted in decreased levels of emotional distress, depression, and anxiety; reduced instances of disruptive, noncompliant, and aggressive behavior; improved attitudes toward self, school, and others; stronger academic motivation and engagement; and an average 11-percentile point increase in academic grades (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). The most efficacious SEL programs utilize a blend of "within-person" strategies, such as explicit lessons that teach students social and emotional competencies directly, and "within-context" strategies, such as instructional practices that create classroom environments conducive to socioemotional learning (Devaney, O'Brien, Resnik, Keister, & Weissberg, 2006). In addition, successful SEL programs build in ample opportunities for students to practice their social and emotional skills both in regular classes (e.g., discussing emotional states of literary characters) and in other school settings (e.g., recess, playground, cafeteria) (CASEL, 2013). Recognizing the important role parents play in children's socialization, many SEL programs also include strategies to engage students' families, in order to extend children's socioemotional learning beyond the school (Albright & Weissberg, 2010). Practices to promote school-family collaboration include educating parents about the benefits of SEL, involving them in setting SEL goals for the school, and providing them with specific strategies to reinforce children's socioemotional competencies at home (Patrikakou, Weissberg, Redding, & Walberg, 2005).

The Collaborative for Academic, Social, and Emotional Learning (CASEL, 2013) has compiled a practitioners' guide to 19 universal school-based SEL programs, selected based on their rigorous multiyear design, availability of implementation supports, and documented positive impact on students' behavior and/or academic performance. This and other resources (e.g., Devaney et al., 2006; Merrell & Gueldner, 2010) can help administrators choose the right SEL program for their district/school. However, the choice of a well-designed SEL program does not in itself guarantee successful outcomes; the extent of the program's impact on students' social, emotional, and academic learning depends on how well it is implemented (Durlak et al., 2011). In turn, high-quality program implementation is

critically dependent on strong support from district/school leaders, who can champion SEL goals within and outside the school, develop infrastructure to support multiyear SEL programming and assessment, provide SEL training and professional development for teachers and staff, and promote systemic integration of SEL practices district/school-wide (CASEL, 2013; Devaney et al., 2006).

Large-scale program dissemination additionally requires buy-in from the various levels of government (Mallin et al., 2013). Given the availability of well-designed SEL programs and the growing research base supporting their efficacy, the SEL framework has seen a steady uptake at the policy level in recent years. In the USA, four states have adopted K-12 educational standards that emphasize social and emotional learning, and most other states have integrated SEL principles into existing standards for other subject areas (Dusenbury, Zadrazil, Mart, & Weissberg, 2011). In addition, the US federal government is currently considering the Academic, Social, and Emotional Learning Act of 2013, which will provide funding for SEL training and professional development of teachers and principals. In Canada, too, many provinces have adopted SEL benchmarks not only for students but also for school leaders and teaching professionals (e.g., British Columbia Ministry of Education, 2013; Ontario Ministry of Education, 2013). These recent policy changes signify an important shift toward greater system-wide effort to promote resilience in North American schools.

12.3.2 Applications at the Classroom Level

Positive classroom climate, caring teacher-student relationships, and stimulating learning environment, these are among the most frequently cited contextual factors that contribute to students' resilience and academic engagement (Song, Doll, & Marth, 2013; Sapienza & Masten, 2011). Importantly, all of these factors are within the educators' control and certainly the mandate of our schools. For this reason, most school-based prevention and intervention programs include a classroom-level component, which involves modifying various contextual and relational influences so as to create conditions that best foster students' sense of relatedness, mastery, and personal control (CASEL, 2013; Song, Sikorski, Doll, & Sikorski, 2014). These contextual and relational influences may range from structured physical spaces, instructional techniques, and classroom management practices to spontaneous teachable moments and teacher attitudes and expectations. To follow is a description of two empirically supported program approaches to promoting resiliency in the classroom: the Responsive Classroom and ClassMaps.

Responsive Classroom A notable example of well-designed, evidence-based classroom practices for enhancing children's resilience and learning is the Responsive Classroom approach developed by the Northeast Foundation for Children (NEFC, 1997). Responsive Classroom practices are based on the principle that children learn best in a safe, caring, and challenging environment that is

responsive to their social, emotional, and intellectual needs. Accordingly, the Responsive Classroom approach aims to alter daily routine, organization, and relational climate of the classroom in a way that enhances both teachers' self-efficacy and children's social and academic performance. For example, the practice of morning meetings is geared toward prosocial skills, community building, and sense of relatedness. Every morning, the class gathers for 30 min to greet one another by name (greeting), share personal news (sharing), and then join in a lively group activity that builds class cohesion and sets a positive tone for the day (group activity). During sharing, students take turns recounting personal news and responding to one another's accounts with questions and comments, while the teacher models and reinforces specific strategies for listening attentively and responding with care and respect. This daily routine is designed to help students practice empathy and communication skills, while fostering a sense of relatedness that comes with being appreciated and understood by one another (Kriete, 2002).

The Responsive Classroom approach to academic instruction and classroom management is proactive rather than reactive and collaborative rather than prescriptive, where students are consistently empowered to take ownership of their own learning and behavior. In academic activities, students are given structured choices over the topics they may pursue, tools they may use, or products they may create (academic choice) which may in turn enhance sense of mastery. In the matters of conduct, too, the teacher actively involves students in creating classroom and activity-specific rules, as well as ways of behaving in accordance with those rules and the consequences for not doing so (rule creation, logical consequences). Prior to undertaking a specific activity, the class is routinely directed to review the relevant rules and to practice the appropriate behaviors, in order to reduce the likelihood of subsequent misbehavior (interactive modeling). These student-centered practices are designed to create a sense of personal agency and responsibility in everything students do at school. In turn, when tasks are seen as intrinsic and meaningful, students are more invested in doing them and putting their best effort forward (Deci, Vallerand, Pelletier, & Ryan, 1991).

The same principles guide the Responsive Classroom approach to problem solving, where the teacher's goal is not to punish or correct misbehavior, but to teach students how to identify and solve their problems autonomously and without losing their dignity (collaborative problem solving). Rather than emphasizing what the student is doing wrong ("Don't do that"), which undermines competence, or telling the student what to do, which takes control away from the student, the teacher may remind the student of the relevant rules and redirect them toward appropriate behaviors ("Show me what you will do to uphold the rules"), thus giving the student an opportunity to demonstrate self-regulation and competence. For more enduring problems, the teacher may engage the student in a one-on-one problem-solving conference. In this technique, the teacher first states his/her observations of the student's behavior in a neutral nonjudgmental tone ("I notice that…") and then invites the student to provide their own thoughts on what is going on ("What do you notice?"), why they think it occurs ("Could it be…?"), and what they might do to resolve it. The student is then encouraged to try one of their own suggested solutions

and to choose a backup solution in case the first one does not work. This collaborative problem-solving approach is designed to teach students how to identify problems, consider different alternatives, and learn from the outcomes, while building a sense of self-efficacy for being able to change their situation for the better (NEFC, 1997; Shure & Aberson, 2013).

An important goal of the Responsive Classroom approach is to encourage teachers to consider the student's perspective and to reexamine their own assumptions about students' behavior. Teachers' beliefs and attributions influence their classroom interactions, which in turn have powerful effects on students' behavior and on the overall classroom climate (Goldstein & Brooks, 2007; McAuliffe, Hubbard, & Romano, 2009; Wiley, Tankersley, & Simms, 2012). Consider, for example, the case of a student, Jonathan, who frequently interrupts the class with endless questions (described by Brooks & Brooks, 2014). The teacher may assume that Jonathan does so deliberately to annoy the teacher, prompting disciplinary action. In contrast, the child's perspective may reveal that the question-asking behavior is driven by high anxiety about not being able to understand the material, combined with hyperactive-impulsive tendencies. In this case, the teacher's misinterpretation of Jonathan's behavior as defiant would not only fail to address his anxiety or meet his learning needs, but it might hurt his academic motivation and convey a negative image of him to his peers. To help teachers avoid unintentionally "punishing a suffering child" and instead become a "charismatic adult" in their students' lives, Brooks and Brooks (2014) encourage educators to ask themselves: "How would I feel if someone said or did to me what I just said or did to my student?", "When I say or do things with my students, am I doing so in a way that will help them realize I love and care about them?", and "Are all of the students in my classroom stronger because of things I've said or done today or are they less strong?". In much the same vein, Responsive Classroom practices (e.g., positive teacher language) are designed to help teachers adopt an empathic attitude toward their students and use nonjudgmental communication style that affirms students' dignity, efforts, and strengths (Denton, 2014).

The efficacy of Responsive Classroom practices has been supported in a series of controlled experiments and longitudinal studies of elementary classrooms (grades K through 4) in schools that have adopted this framework versus schools that did not. Teachers' use of Responsive Classroom practices was associated with improved quality of teacher-student interactions, more favorable student perceptions of the school, enhanced social skills, and improved performance on standardized reading and math tests (Abry, Rimm-Kaufman, Larsen, & Brewer, 2013; Brock, Nishida, Chiong, Grimm, & Rimm-Kaurman, 2008; Rimm-Kaufman & Chiu, 2007; Rimm-Kaufman, Fan, Chiu, & You, 2007). In addition to gains in student outcomes, the Responsive Classroom approach also contributed to greater teacher effectiveness. Teachers who used Responsive Classroom practices provided more emotional support to their students, held more positive attitudes toward their own teaching and the teaching profession in general, and had stronger commitment to promoting students' active learning, prosocial skills, and self-control (Curby, Rimm-Kaufman, & Abry, 2013; Rimm-Kaufman & Sawyer, 2004).

The impact of Responsive Classroom practices is stronger when they are applied school-wide (Wanless, Patton, Rimm-Kaufman, & Deutsch, 2013). However, an important advantage of this and other classroom-level interventions is that they can be implemented as stand-alone initiatives as well, carried out by individual teachers. This makes them particularly attractive for schools that do not have the requisite resources to roll out and maintain comprehensive SEL-type programming (Embry & Biglan, 2008). In either case, it is advisable that teachers receive adequate training in the appropriate skills and strategies, to maximize their capacity for nurturing resiliency in their students (CASEL, 2013; Song et al., 2014; Wanless et al., 2013).

ClassMaps ClassMaps, developed by Doll, Brehm, and Zucker (2014), is a data-based consultation model to help teachers modify ordinary classroom environments so that these are more resilience promoting. The strategy promotes resilience by assessing the classroom characteristics that enhance the developing child's relationships or support the child's emerging human agency. Subscales of the ClassMaps assessment represent several aspects of interactive resilience. Five subscales describe relational aspects of the classroom, including teacher-student relationships, peer friendships, peer conflict, worries about peer aggression, and home-school relationships. Three of the subscales describe autonomy characteristics, including academic self-efficacy, self-determination, and behavioral self-control. The underlying assumption of the ClassMaps approach is that helping teachers change their classrooms to be experienced as more resilience supporting will enhance the resiliency and learning of the students in the classroom.

Assessment data collected from students are used to identify limitations in the classroom's interpersonal relationships or its routines and practices that undermine student autonomy. Based on this needs assessment and drawing from the recommendations of students, classroom teachers can identify the aspects of the classroom that are the best targets for intervention. Then, because both teachers and students are highly familiar with classroom routines, they can propose changes that are likely to strengthen the classroom's relational and autonomy characteristics. Subsequently, classroom data can be used to monitor the impact that these changes have had and guide teachers' decisions to continue, intensify, or alter their plans for change.

An important focus of the ClassMaps Consultation research was the identification of a brief and technically sound assessment of these resilience-promoting characteristics of classrooms (Doll et al., 2014). ClassMaps is a 55-item anonymous student survey with eight subscales that are aggregated across students in the surveyed classroom. Three subscales assess the collective self-regulation of students in the class: academic efficacy (believing in me), academic self-determination (taking charge), and behavioral self-control (following class rules). Five subscales assess the classroom relationships: teacher-student relationships (my teacher), home-school relationships (talking with my parents), peer friendships (my classmates), peer conflict (kids in this class), and concerns about bullying (I worry that). Students select their response from a four-point scale (never, sometimes, often, almost always). When computer administered, the survey is completed by the entire class in about 15–20 min. The internal consistency, factor structure, and concurrent validity

of the full ClassMaps Survey and its subscales have been examined with elementary students (Doll, Spies, LeClair, Kurien, & Foley, 2010) and secondary students (Doll, Spies, Champion et al., 2010).

Teachers use data from the ClassMaps Survey to examine students' perceptions of the classroom resilience, and then they discuss the data with colleagues and/or the classrooms' students to check the accuracy of their understanding. Incidentally, sharing the classroom data provides teachers with alternative strategies for change, builds a support system for change, and fosters the students' ownership for classroom change. In response to identified weaknesses in the classroom's protective factors, teachers may elect to make microchanges (informal adjustments to classroom routines) or to implement manualized classroom interventions (which have been developed and examined in well-controlled classroom intervention research). Because the teachers are collecting data to monitor the effects of the interventions, they will know if a microchange was insufficient (and can transition to a more intensive manualized intervention) or if a manualized intervention was incompatible with the culture of the classroom (and can transition to a modified intervention that is sensitive to local conditions).

To date, evidence for the impact of ClassMaps Consultation has occurred through small-n research (Murphy, 2002; Nickolite & Doll, 2008) and case studies (Doll et al., 2014). With the availability of a new data curriculum, it will be possible to examine the impact of the procedure with larger groups of teachers using better controlled experimental designs. Ultimately, ClassMaps Consultation has the potential to embed protective factors into daily classroom environments, taking advantage of teachers' familiarity with their students' development and their expertise in classroom systems. The goal is to provide teachers with a process for classroom improvement that builds on the compelling research on developmental resiliency and infuses strengthened protective factors into school environments.

12.3.3 Applications at the Individual Level

Although safe and responsive school environments are vitally important in promoting positive development for all children, most resilience scholars agree that the effects of external protective factors are both mediated and moderated by the individual's subjective experiences (Brock et al., 2008; Brooks & Brooks, 2014; Masten, 2001; Prince-Embury, 2013b; Song et al., 2014). Indeed, resilient outcomes are often more strongly related to one's perceived competence rather than actual abilities (Bandura, 1993), perceived availability of social support rather than actual supports available (Cohen & Wills, 1985), and perceived degree of control over outcomes rather than actual event controllability (Lazarus & Folkman, 1984). From a practical standpoint, this implies that even when the classroom conditions are optimally conducive to promoting resiliency, individual students may need further intervention to help them internalize these conditions into a subjective sense of

self-efficacy, relatedness, adaptability, and personal control – qualities that make up a "resilient mindset" (Brooks & Goldstein, 2001).

In describing ways to foster resilient mindsets in individual children, Brooks, Goldstein, and colleagues (Brooks & Brooks, 2014; Brooks, Brooks, & Goldstein, 2012; Goldstein, Brooks, & DeVries, 2013) emphasize the importance of cultivating "islands of competence," or areas of personal strength, in every child. When individuals discover that they can be successful at something, particularly in an area that is important to them and valued by significant others, they are more likely to draw on that strength for global feelings of self-efficacy and self-worth, which subsequently spread to other areas of self-concept (McConnell, 2011). Accordingly, Brooks, Goldstein, and colleagues encourage teachers to make a list of their students' individual strengths and competencies and to come up with ways to reinforce those islands of competence in their everyday interactions with the students.

When attempting to change students' mindsets, it is important to understand the psychological mechanisms involved in the formation of self-referent cognitions. Although competence self-perceptions derive from multiple sources, personal mastery experiences exert by far the most powerful influence on self-efficacy beliefs (Bandura, 1993; Usher & Pajares, 2008). When students repeatedly achieve successful outcomes, their sense of competence is strengthened and so is their confidence in doing well in the future. In contrast, when students repeatedly fail in their attempts to achieve the desired outcomes, they begin to doubt their abilities and lose hope that things will change for the better. The resulting mindsets both contribute to and are perpetuated by subsequent experiences in a reciprocal fashion. Students who have developed a self-efficacious mindset are more likely to seek challenges, try harder, persevere in the face of setbacks, and ultimately fulfill their goals (Bandura, 1993). In contrast, equally able students who have come to believe that they have no control over their outcomes (i.e., helpless mindset) tend to avoid challenges, put forth less effort, give up after setbacks, and as a result are less likely to discover that they can affect positive change in their lives (Seligman, 1990).

In practice, this means that providing students with opportunities to experience success directly is the most effective and authentic way to build up their sense of mastery, competence, and personal control (Goldstein et al., 2013). An important caveat to remember here is that mastery experiences are inherently subjective, for the same level of performance may be interpreted as a success by one student and a failure by another (Usher & Pajares, 2008). These interpretations depend on a series of temporal, dimensional, and social comparisons students make in relation to their past performance, their performance in other areas, and performance of their peers (Möller, 2005). Students are more likely to experience increases in self-efficacy when their performance improves over time, when it is in the domain that is important to them, and when they are doing better than their classmates (Marsh, 2007). Students' interpretations of their performance also depend on adults' expectations for them: unrealistically high expectations set students up for failure regardless of how capable they are, whereas very low expectations trivialize students' success by

implying they are not capable of doing better (Goldstein et al., 2013). Indeed, the most powerful mastery experiences occur when students successfully overcome obstacles or accomplish challenging but manageable tasks (Bandura, 1993).

The attributions students make about causes of their successes and failures represent another source of individual differences in resilient mindsets (Brooks & Brooks, 2014). When students attribute their performance outcomes to factors that are within their personal control (e.g., effort, resources), they are more likely to internalize successes and treat failures as temporary setbacks and opportunities to learn. In contrast, when students attribute their performance outcomes to factors they cannot change (e.g., genes, luck), they are not only less likely to benefit from positive mastery experiences but also more likely to adopt self-defeating ways of coping with failure, such as disengagement, self-handicapping, or blaming others. Children acquire these attributional styles through vicarious observations and through explanations of successes and failures provided to them by significant others (Bandura, 1993; Frome & Eccles, 1988). Thus, teachers may become influential sources of resilient ways of thinking for their students, by modeling constructive ways of dealing with challenges and by emphasizing situations where a student's effort had a direct impact on the outcome (Brooks & Brooks, 2014; Goldstein et al., 2013).

Another salient source of students' self-efficacy beliefs is evaluative feedback received from significant others (Bandura, 1993; Usher & Pajares, 2008). At the very least, teachers may provide verbal encouragement to their students, communicating that they notice and value students' strengths and not just focus on their weaknesses. However, teachers also need to be aware that their verbal affirmations may not always appear welcome. Indeed, the very individuals who would benefit from encouragement the most, i.e., those with low self-esteem and little self-confidence, are often the least receptive to positive feedback, dismissing it as fundamentally incongruent with their sense of who they are (Swann, 1997). Likewise, individuals suffering from depression tend to be less accepting of others' expressions of love and support for them, which often elicits frustration and rejection by others and thereby reinforces the depressive mindset pervaded by feelings of shame and worthlessness (Joiner, Katz, & Lew, 1997). Regardless of individual circumstances, it is important to continue verbally affirming students' strengths, but at the same time recognizing that changing such negative mindsets may require more intensive cognitive-behavioral intervention (Goldstein et al., 2013).

An important take-home message for teachers is that while mastery experiences, effort attributions, and positive feedback may not always increase students' self-efficacy, repeated experiences of failure, attributions to lack of ability, and predominantly negative feedback are almost certain to erode students' self-worth (Usher & Pajares, 2008). It is for this reason that Brooks, Goldstein, and colleagues (Brooks & Brooks, 2014; Brooks et al., 2012; Goldstein et al., 2013) encourage teachers to focus their interventions not only on the areas where students are struggling but also on those islands of competence where students are already doing well and that matter to them a great deal. Once a foundation for mastery is established, a more resilient mindset will follow.

12.4 Summary

This chapter has discussed "resiliency" in the context of current and future applications in school settings. The first part of the chapter described the constructs of resilience/resiliency and presented a three-factor model of personal resiliency which simplifies the construct into three developmental systems that may in turn be used for focused application in schools. The Resiliency Scales for Children and Adolescents were then presented as a tool for preventive screening in schools that employ the three-factor model of personal resiliency. The second part of the chapter presented interventions, currently in use, that employ strength-based principles including aspects of resilience/resiliency. We acknowledge that the application of strength- and resilience-based models into the education system is an uneven process characterized by a different terminology, focus, technique, and level of application. We hope however that these differences do not impede the important process of integrating the science of resilience and strength building into an important pre-existing infrastructure for the education and development of our children.

References

Abry, T., Rimm-Kaufman, S. E., Larsen, R. A., & Brewer, A. J. (2013). The influence of fidelity of implementation on teacher-student interaction quality in the context of a randomized controlled trial of the Responsive Classroom approach. *Journal of School Psychology, 51*, 437–453.

Albright, M. I., & Weissberg, R. P. (2010). School-family partnerships to promote social and emotional learning. In S. L. Christenson & A. L. Reschley (Eds.), *The handbook of school-family partnerships for promoting student competence* (pp. 246–265). New York: Routledge.

Alvord, M. K., Zucker, B., & Grados, J. J. (2011). *Resilience builder program: Enhancing social competence and self-regulation*. Champaign, IL: Research Press.

Bandura, A. (1977). *Self-efficacy: The exercise of control*. New York: Freeman.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117–148.

Bar-On, R. (2006). The Bar-On model of emotional-social intelligence. *Psicothema, 18*, 13–25.

Block, J. H., & Block, J. (1980). The role of ego-control and ego-resiliency in the origination of behavior. In W. A. Collings (Ed.), The Minnesota Symposia on Child Psychology (Vol. 13, pp. 39–101). Hillsdale, NJ: Erlbaum.

Block, J. (2002). Personality as an affect processing system. Mahwah, N.J. Erlbaum.

British Columbia Ministry of Education. (2013). *Defining cross-curricular competencies: Transforming curriculum and assessment*. Victoria: British Columbia Ministry of Education.

Brock, L. L., Nishida, K. K., Chiong, C., Grimm, K. J., & Rimm-Kaurman, S. E. (2008). Children's perceptions of the social environment and social and academic performance: A longitudinal analysis of the Responsive Classroom approach. *Journal of School Psychology, 46*, 129–149.

Brooks, R., & Brooks, S. (2014). Creating resilient mindsets in children and adolescents: A strength-based approach for clinical and nonclinical populations. In S. Prince-Embury & D. H. Saklofske (Eds.), *Resilience interventions for youth in diverse populations* (pp. 59–82). New York: Springer.

Brooks, R., Brooks, S., & Goldstein, S. (2012). The power of mindsets: Nurturing engagement, motivation, and resilience in students. In S. Christenson, A. L. Resschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 541–562). New York: Springer.

Brooks, R., & Goldstein, S. (2001). *Raising resilient children*. New York: McGraw-Hill.

Brooks, R., & Goldstein, S. (2008). The mindset of teachers capable of fostering resilience in students. *Canadian Journal of School Psychology, 23*(1), 114–126.

Cicchetti, D., Ganiban, J., & Barnett, D. (1991). Contributions from the study of high-risk populations to understanding the development of emotion regulation. In J. Garber & K. Dodge (Eds.), *The development of emotion regulation and dysregulation* (pp. 15–48). New York: Cambridge University Press.

Cicchetti, D., & Tucker, D. (1994). Development and self-regulatory structures of the mind. *Development and Psychopathology, 6*(4), 533–549.

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin, 98*, 310–357.

Collaborative for Social and Emotional Learning. (2013). *2013 CASEL guide: Effective social and emotional learning programs – Preschool and elementary school edition*. Chicago, IL: Collaborative for Social and Emotional Learning.

Curby, T. W., Rimm-Kaufman, S. E., & Abry, T. (2013). Do emotional support and classroom organization earlier in the year set the stage for higher quality instruction? *Journal of School Psychology, 51*, 557–569.

de Boo, G. M., & Prins, P. J. M. (2007). Social incompetence in children with ADHD: Possible moderators and mediators in social-skills training. *Clinical Psychology Review, 27*(1), 78–97.

Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist, 26*, 325–346.

Denton, P. (2014). *The power of our words: Teacher language that helps children learn* (2nd ed.). Turners Falls, MA: Northeast Foundation for Children.

Devaney, E., O'Brien, M. U., Resnik, H., Keister, S., & Weissberg, P. (2006). *Sustainable schoolwide social and emotional learning (SEL): Implementation guide and toolkit*. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning.

Doll, B., Brehm, K., & Zucker, S. (2014). *Resilient classrooms: Creating healthy environments for learning* (2nd ed.). New York: Guilford Press.

Doll, B., Spies, R. A., Champion, A., Guerrero, C., Dooley, K., & Turner, A. (2010). The ClassMaps Survey: A measure of students' perceptions of classroom resilience. *Journal of Psychoeducational Assessment, 28*, 338–348.

Doll, B., Spies, R. A., LeClair, C., Kurien, S., & Foley, B. (2010). Student perceptions of classroom learning environments: Development of the ClassMaps Survey. *School Psychology Review, 39*, 203–218.

Doll, B., Zucker, S., & Brehm, K. (2004). *Resilient classrooms: Creating healthier environments for learning*. New York: Guilford Press.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*, 405–432.

Dusenbury, L., Zadrazil, J., Mart, A., & Weissberg, R. P. (2011). *State learning standards to advance social and emotional learning: The state scan of social and emotional learning standards, preschool through high school*. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning.

Eisenberg, N., Champion, C., & Ma, Y. (2004). Emotion-related regulation: An emerging construct. *Merrill-Palmer Quarterly, 50*, 236–259.

Elias, M. J., Parker, S. R., & Rosenblatt, J. L. (2005). Building educational opportunity. In R. B. Brooks & S. Goldstein (Eds.), *Handbook of resilience in children* (pp. 315–336). New York: Kluwer Academic/Plenum.

Embry, D. D., & Biglan, A. (2008). Evidence-based kernels: Fundamental units of behavioral influence. *Clinical Child and Family Psychology Review, 11*, 75–113.

Frome, P. M., & Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of Personality and Social Psychology, 74*, 435–452.

Garmezy, N. (1971). Vulnerability research and the issue of primary prevention. *American Journal of Orthopsychiatry, 41*, 101–116.

Garmezy, N. (1985). Stress-resistant children: The search for protective factors. In J. E. Stevenson (Ed.), *Recent research in developmental psychopathology (Journal of Child Psychology and Psychiatry Book Suppl. 4*, pp. 213–233). Oxford: Pergamon.

Garmezy, N. (1991). Resilience and vulnerability to adverse developmental outcomes associated with poverty. *American Behavioral Scientist, 34*, 416–430.

Garmezy, N., Masten, A. S., & Tellegen, A. (1984). The study of stress and competence in children: A building block for developmental psychopathology. *Child Development, 55*, 97–111.

Goldstein, S., & Brooks, R. (2007). *Understanding and managing classroom behavior: Creating resilient, sustainable classrooms*. New York: Wiley.

Goldstein, S., Brooks, R., & DeVries, M. (2013). Translating resilience theory for application with children and adolescents by parents, teachers, and mental health professionals. In S. Prince-Embury & D. H. Saklofske (Eds.), *Resilience in children, adolescents, and adults: Translating research into practice* (pp. 73–90). New York: Springer.

Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., et al. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist, 58*, 466–474.

Gresham, F. M., & Elliot, S. N. (1990). *Social skills rating system manual*. Circle Pines, MN: American Guidance Service.

Joiner, T. E., Jr., Katz, J., & Lew, A. S. (1997). Self-verification and depression among youth psychiatric inpatients. *Journal of Abnormal Psychology, 106*, 608–618.

Kaplan, H. B. (1999). Toward an understanding of resilience: A critical review of definitions and models. In M. D. Glantz & J. L. Johnson (Eds.), *Resilience and development: Positive life adaptations* (pp. 17–83). New York: Kluwer Academic/Plenum.

Kaplan, H. B. (2005). Understanding the concept of resilience. In S. Goldstein & R. Brooks (Eds.), *Handbook of resilience in children* (pp. 39–47). New York: Springer.

Kriete, R. (2002). *The morning meeting book*. Turners Falls, MA: Northeast Foundation for Children.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.

Luthar, S. S. (1991). Vulnerability and resilience: A study of high-risk adolescents. *Child Development, 62*, 600–616.

Luthar, S. S. (2006). Resilience in development: A synthesis of research across five decades. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Risk, disorder, and adaptation* (2nd ed., Vol. 3, pp. 739–795). Hoboken, NJ: Wiley.

Luthar, S. S., Cicchetti, D. C., & Becker, B. (2000). The construct of resilience: A critical evaluation and guidelines for future work. *Child Development, 71*, 543–562.

Luthar, S. S., & Zelazo, L. B. (2003). Research on resilience: An integrative review. In S. S. Luthar (Ed.), *Resilience and vulnerability: Adaptation in the context of childhood adversities* (pp. 510–549). New York: Cambridge University Press.

Luthar, S. S., & Zigler, E. (1991). Vulnerability and competence: A review of the research on resilience in childhood. *American Journal of Orthopsychiatry, 61*, 6–22.

Luthar, S. S., & Zigler, E. (1992). Intelligence and social competence among high-risk adolescents. *Development and Psychopathology, 4*, 287–299.

Malecki, C. K., & Elliott, S. N. (2002). Children's social behaviors as predictors of academia achievement: A longitudinal analysis. *School Psychology Quarterly, 17*, 1–23.

Mallin, B., Walker, J. R., & Levin, B. (2013). Mental health promotion in the schools: Supporting resilience in children and youth. In S. Prince-Embury & D. H. Saklofske (Eds.), *Resilience in children, adolescents, and adults: Translating research into practice* (pp. 91–112). New York: Springer.

Marsh, H. W. (2007). *Self-concept theory, measurement and research into practice: The role of self concept in educational psychology*. Leicester: British Psychological Society.

Masten, A. S. (2001). Ordinary magic: Resilience processes in development. *American Psychologist, 56*, 227–238.

Masten, A. S., & Coatsworth, J. D. (1998). The development of competence in favorable and unfavorable environments: Lessons from research on successful children. *American Psychologist, 53*, 205–220.

Masten, A. S., & Curtis, W. J. (2000). Integrating competence and psychopathology: Pathways toward a comprehensive science of adaptation in development [Special issue]. *Development & Psychopathology, 12*, 529–550.

Masten, A. S., & Obradovic, J. (2006). Competence and resilience in development. *Annals of the New York Academy of Sciences, 1094*, 13–22.

Masten, A. S., & Powell, J. L. (2003). A resilience framework for research, policy, and practice. In S. S. Luthar (Ed.), *Resilience and vulnerability: Adaptation in the context of childhood adversities* (pp. 1–25). New York: Cambridge University Press.

Masten, A. S., Roisman, G. I., Long, J. D., Burt, K. B., Obradovic, J., Riley, J. R., et al. (2005). Developmental cascades: Linking academic achievement and externalizing and internalizing symptoms over 20 years. *Developmental Psychology, 41*, 733–746.

McAuliffe, M. D., Hubbard, J. A., & Romano, L. J. (2009). The role of teacher cognition and behavior in children's peer relations. *Journal of Abnormal Child Psychology, 37*, 665–677.

McConnell, A. R. (2011). The multiple self-aspects framework: Self-concept representation and its implications. *Personality and Social Psychology Review, 15*, 3–27.

McDermott, P. A., Mordell, M., & Stoltzfus, J. (2001). The organization of student performance in American schools: Discipline, motivation, verbal learning, and nonverbal learning. *Journal of Educational Psychology, 93*, 65–76.

Merrell, K. W., & Gueldner, B. A. (2010). *Social and emotional learning in the classroom: Promoting mental health and academic success*. New York: The Guilford Press.

Möller, J. (2005). Paradoxical effects of praise and criticism: Social, dimensional and temporal comparisons. *British Journal of Educational Psychology, 75*, 275–295.

Murphy, P. S. (2002). *The effect of classroom meetings on the reduction of recess problems: A single case design.* Unpublished doctoral dissertation, University of Denver, Denver, CO.

Nickolite, A., & Doll, B. (2008). Resilience applied in school: Strengthening classroom environments for learning. *Canadian Journal of School Psychology, 23*, 94–113.

Northeast Foundation for Children (NEFC). (1997). *Guidelines for the responsive classroom*. Greenfield, MA: Northeast Foundation for Children.

Ontario Ministry of Education. (2013). *Ontario leadership strategy: Principal/vice-principal performance appraisal – technical requirements manual*. Toronto, ON: Ontario Ministry of Education.

Patrikakou, E. N., Weissberg, R. P., Redding, S., & Walberg, H. J. (Eds.). (2005). *School-family partnerships for children's success*. New York: Teachers College Press.

Petrides, K. V. (2011). Ability and trait emotional intelligence. In T. Chamorro-Premuzic, S. von Stumm, & A. Furnham (Eds.), *The Wiley-Blackwell handbook of individual differences* (pp. 656–678). Oxford: Wiley-Blackwell.

Prince-Embury, S. (2006) The Resiliency Scales for Adolescents, Harcourt Assessment. San Antonio, Tx.

Prince-Embury, S. (2007). *Resiliency scales for children and adolescents: Profiles of personal strengths*. San Antonio, TX: Harcourt Assessments.

Prince-Embury, S. (2008). Resiliency scales for children and adolescents, psychological symptoms and clinical status of adolescents. *Canadian Journal of School Psychology, 23*, 41–56.

Prince-Embury, S. (2010). Assessment for integrated screening and prevention using the resiliency scales for children and adolescents. In B. Doll, W. Pfohl, & J. Yoon (Eds.), *Handbook of youth prevention science* (pp. 141–162). New York: Routledge.

Prince-Embury, S. (2011). Assessing personal resiliency in the schools using the resiliency scales for children and adolescents. *Psychology in the Schools, 48*(7), 672–685.

Prince-Embury, S. (2013a). Translating resilience theory for assessment and application with children, adolescents, and adults: Conceptual issues. In S. Prince-Embury & D. H. Saklofske (Eds.), *Resilience in children, adolescents, and adults: Translating research into practice* (pp. 9–16). New York: Springer.

Prince-Embury, S. (2013b). Resiliency scales for children and adolescents: Theory, research, and clinical application. In S. Prince-Embury & D. H. Saklofske (Eds.), *Resilience in children, adolescents, and adults: Translating research into practice* (pp. 19–44). New York: Springer.

Prince-Embury, S. (2014). Three factor model of personal resiliency and related interventions. In S. Prince-Embury & D. H. Saklofske (Eds.), *Resilience interventions for youth in diverse populations* (pp. 25–57). New York: Springer.

Prince-Embury, S., & Courville, T. (2008a). Comparison of one, two and three factor models of personal resilience using the resiliency scales for children and adolescents. *Canadian Journal of School Psychology, 23*, 11–25.

Prince-Embury, S., & Courville, T. (2008b). Measurement invariance of the resiliency scales for children and adolescents with respect to sex and age cohorts. *Canadian Journal of School Psychology, 23*, 26–40.

Prince-Embury, S., & Saklofske, D. H. (2013). Translating resilience theory for application: Introduction. In S. Prince-Embury & D. H. Saklofske (Eds.), *Resilience in children, adolescents, and adults: Translating research into practice* (pp. 3–7). New York: Springer.

Prince-Embury, S., & Saklofske, D. H. (2014). Building a science of resilience intervention for youth. In S. Prince-Embury & D. H. Saklofske (Eds.), *Resilience interventions for youth in diverse populations* (pp. 3–12). New York: Springer.

Rimm-Kaufman, S. E., & Chiu, Y. I. (2007). Promoting social and academic competence in the classroom: An intervention study examining the contribution of the Responsive Classroom approach. *Psychology in the Schools, 44*, 397–413.

Rimm-Kaufman, S. E., Fan, X., Chiu, Y. I., & You, W. (2007). The contribution of the Responsive Classroom approach on children's academic achievement: Results of a three-year longitudinal study. *Journal of School Psychology, 45*, 401–421.

Rimm-Kaufman, S. E., & Sawyer, B. E. (2004). Primary grade teachers' self-efficacy beliefs, attitudes toward teaching, and discipline and teaching practice priorities in relation to the Responsive Classroom approach. *The Elementary School Journal, 104*, 321–341.

Rothbart, M. K., & Derryberry, D. (1981). Development of individual differences in temperament. In M. E. Lamb & A. L. Brown (Eds.), *Advances in developmental psychology* (Vol. 1, pp. 37–86). Hillsdale, NJ: Erlbaum.

Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *American Journal of Orthopsychiatry, 57*, 316–331.

Rutter, M. (1993). Resilience: Some conceptual considerations. *Journal of Adolescent Health, 14*, 626–631.

Rutter, M. (2010). Child and adolescent psychiatry: Past scientific achievement and challenges for the future. *European Child and Adolescent Psychiatry, 19*, 689–703. doi:10.1007/s00787-010-0111-y.

Sapienza, J. K., & Masten, A. S. (2011). Understanding and promoting resilience in children and youth. *Current Opinion in Psychiatry, 24*, 267–273.

Schwean, V. L., & Rodger, S. (2013). Children first: It's time to change! Mental health promotion, prevention, & treatment informed by public health and resiliency approaches. *Canadian Journal of School Psychology, 28*, 136–166.

Seligman, M. E. P. (1990). *Learned optimism: How to change your minds and life*. New York: Pocket Books.

Shure, M. B., & Aberson, B. (2013). Enhancing the process of resilience through effective thinking. In S. Goldstein & R. Brooks (Eds.), *Handbook of resilience in children* (2nd ed., pp. 459–480). New York: Springer.

Siegel, D. J. (1999). *The developing mind: How relationships and the brain interact to shape who we are*. New York: Guilford Press.

Song, S. Y., Doll, B., & Marth, K. (2013). Classroom resilience: Practical assessment for intervention. In S. Prince-Embury & D. H. Saklofske (Eds.), *Resilience in children, adolescents, and adults: Translating research into practice* (pp. 61–72). New York: Springer.

Song, S. Y., Sikorski, J., Doll, B., & Sikorski, M. (2014). Enhancing classroom resilience with ClassMaps consultation. In S. Prince-Embury & D. H. Saklofske (Eds.), *Resilience interventions for youth in diverse populations* (pp. 203–215). New York: Springer.

Swann, W. B. (1997). The trouble with change: Self-verification and allegiance to the self. *Psychological Science, 8*, 177–180.

Thompson, R. A. (1990). Emotion and self-regulation. In R. Dienstbier (Series Ed.) & R. A. Thompson (Vol. Ed.), *Nebraska symposium on motivation: Socioemotional development* (pp. 367–467). Lincoln: University of Nebraska Press.

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*, 751–796.

Wang, M. (1994). *Educational resilience in inner-city America: Challenges and prospects*. Hillsdale, NJ: Erlbaum.

Wanless, S. B., Patton, C. S., Rimm-Kaufman, S. E., & Deutsch, N. L. (2013). Setting-level influences on implementation of the Responsive Classroom approach. *Prevention Science, 14*, 40–51.

Wentzel, K. R., & Watkins, D. E. (2002). Peer relationships and collaborative learning as contexts for academic enablers. *School Psychology Review, 31*, 366–377.

Werner, E. E., & Smith, R. S. (1982). *Vulnerable but invincible: A longitudinal study of resilient children and youth*. New York: McGraw-Hill.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66*, 297–333.

Wiley, A. L., Tankersley, M., & Simms, A. (2012). Teachers' causal attributions for student problem behavior: Implications for school-based behavioral interventions and research. In B. G. Cook, M. Tankersley, & T. J. Landrum (Eds.), *Classroom behavior, contexts, and interventions* (pp. 279–300). Bingley, UK: Emerald Group Publishing Limited.

Wyman, P. A., Cowen, E. L., Work, W. C., & Kerley, J. H. (1993). The role of children's future expectations in self-system functioning and adjustment to life stress: A prospective study of urban at-risk children (Special issue). *Development and Psychopathology, 5*, 649–661.

Zins, J. E., Bloodworth, M. R., Weissberg, R. P., & Walberg, H. J. (2004). The scientific base linking social and emotional learning to school success. In J. E. Zins, R. P. Weissberg, M. C. Wang, & H. J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (pp. 3–22). New York: Teachers College Press.

Chapter 13
Teaching Emotional Intelligence in Schools: An Evidence-Based Approach

Catalina Torrente, Susan E. Rivers, and Marc A. Brackett

13.1 Introduction

The skills of emotional intelligence such as recognizing emotions in the self and others, understanding the causes and consequence of emotions, and regulating emotions are fundamental for children's academic and social success (e.g., Blair, Denham, Kochanoff, & Whipple, 2004; Davis & Levine, 2013; Eggum et al., 2011; Eisenberg et al., 1997; Gumora & Arsenio, 2002; Rosen, Milich, & Harris, 2012; Trentacosta & Izard, 2007). These skills enable children to learn and form supportive relationships with adults and peers and are the basis of mental health. However, many children do not arrive at school with the skills they need to manage both the academic and social demands of school and attain positive outcomes (e.g., Rimm-Kaufman, Pianta, & Cox, 2000). For this reason, the premise that schools ought to have a more deliberate role in promoting children's social and emotional skills recently has gained traction among educational leaders, researchers, and policy-makers (Sherman, 2011).

As schools strive to support students' social and emotional development, demand for evidence-based social and emotional learning (SEL) approaches for teaching these skills has increased, along with the need for instruments to both assess related skills and evaluate the effectiveness of programs in improving student outcomes. In this chapter, we provide an overview of RULER, a CASEL SELect approach to SEL developed at the Yale Center for Emotional Intelligence, and elaborate on its

C. Torrente, Ph.D.
Mathematica Public Policy, Yale Center for Emotional Intelligence, Yale University,
340 Edwards Street, New Haven, CT 06520-8376, USA
e-mail: catalina.torrente@gmail.com

S.E. Rivers, Ph.D. (✉) • M.A. Brackett, Ph.D.
Yale Center for Emotional Intelligence, Yale University,
340 Edwards Street, New Haven, CT 06520-8376, USA
e-mail: susan.rivers@yale.edu; marc.brackett@yale.edu

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_13

theoretical underpinnings and evidence base. We also discuss how measurement of children's skills can be strategically used to improve teaching practices and SEL program design, implementation, and effectiveness.

RULER teaches a set of skills collectively referred to as emotional intelligence: recognizing emotions in the self and others, understanding the causes and consequences of emotion, labeling emotions using a sophisticated vocabulary, expressing emotions in socially appropriate ways, and regulating emotions effectively to meet personal and social goals. Emotion regulation stands out as one of the most critical skills for children's functioning in and out of school (Blair, 2002). Generally, children with more developed emotion regulation skills perform better academically and socially and are less likely to develop behavioral and mental health problems (Calkins & Keane, 2004; Trentacosta & Izard, 2007). Given the significance of emotion regulation and limited space to describe all aspects of RULER, we present a brief description of RULER's instructional strategies and tools for teaching emotion regulation and ensuring the skills are embedded into the routines and ethos of schools.

We begin the chapter with an explanation of the role of emotions in children's school experiences, with a particular focus on the development of emotion regulation in childhood. Examining a wide range of emotion regulation strategies, including how they develop in early to late childhood, provides a context for understanding best practices in the design and evaluation of SEL programs. We then present RULER's theory of change to demonstrate how one theoretically driven SEL program addresses the development of emotion regulation in children. Next, we illustrate how RULER integrates the teaching of emotion regulation skills into the core academic curricula. We conclude by discussing how the formative assessment of emotion regulation can be used to fine-tune SEL interventions to match students' needs and to identify areas of success and in need of improvement.

13.1.1 Emotions Matter

Emotions are an inherent component of all our experiences, from childhood through adulthood (see Chap. 11, for review). Even when we are not aware of our emotions, our emotional system is continuously monitoring the environment for changes that may be relevant to our goals, values, and well-being (Cole, Martin, & Dennis, 2004). As such, emotions underlie every experience and direct our behavioral responses to those experiences (Davis & Levine, 2013; Thompson & Meyer, 2007). Emotions can influence the self and others by indicating the need to initiate, maintain, or modify a particular course of action (e.g., Sutton & Wheatley, 2003). Emotions can also organize, facilitate, or hinder internal psychological processes, such as the deployment of attention, decision-making, and the ability to solve problems, which are critical for academic performance (Fischer, Shaver, & Carnochan, 1990). For example, when an injustice is perceived, such as when an aggressor takes away our lunch or pushes us, we are likely to experience anger. Anger directs our attention to the injustice; prepares our body for action by changing our heart rate,

posture, and muscle tension; and motivates us to behave in ways conducive to resisting or overcoming the event, such as grabbing our money back or asking an adult for help. Because attention is focused on the precipitating event, we may forget temporarily other important goals, such as avoiding harm or getting to our next class (Fischer et al., 1990).

Despite its important role, the emotion system is imperfect. In order to efficiently appraise the significance of a situation, emotions unfold very rapidly and may lead to responses that are not always in the individual's best interest. For that reason, children need to learn effective strategies to regulate the myriad emotions that are part of their normal school experiences, such as feeling anxious about tests, ecstatic about winning a race, excluded by a group of peers, nervous about receiving feedback, or enthusiastic about a field trip. In addition, children (and adults) arrive in the classroom differentially prepared to understand and manage emotional information (Raver, 2002). Some children need to learn to effectively regulate their emotions before they can begin to engage with the academic and social opportunities afforded by the classroom. Children who arrive at school being overly aggressive, inattentive, socially anxious, or withdrawn may miss early opportunities to build academic and social skills that lay the stage for later achievements, thus setting in motion a chain of events conducive to academic and social failure (Shields, Cicchetti, & Dyan, 1994).

13.1.2 Emotion Regulation

Emotion regulation is a skill that educators consider central for children's adaptation to school (Blair, 2002; Rimm-Kaufman et al., 2000). Children who are able to regulate their emotions are more likely to pay attention during class, control their anxiety during exams, and make and maintain friendships (Bradley et al., 2010; Fabes et al., 1999; Graziano, Reavis, Keane, & Calkins, 2007). Emotion regulation has direct and indirect effects on children's school adaptation. Effective emotion regulation directly impacts learning by enabling children to use higher-order cognitive processes such as attention, working memory, and planning, all of which are fundamental for mastering new skills and acquiring new knowledge (Blair, 2002). Emotion regulation can contribute indirectly to early social and academic success by allowing children to adjust to the behavioral norms of the classroom and to engage in more harmonic and supportive relationships with teachers and peers (Graziano et al., 2007). For example, children with better regulation skills may be more likely to elicit warm and positive relationships with teachers. On the contrary, children who have difficulties regulating their emotions and behaviors may elicit more conflict in social interactions (e.g., Jerome, Hamre, & Pianta, 2009; Koles, O'Connor, & McCartney, 2009; Spilt, Hughes, Wu, & Kwok, 2012).

Even though there is consensus about the importance of emotion regulation, agreeing on a working definition has been challenging (for a description of these challenges, see Cole et al., 2004; Esbjørn, Bender, Reinholdt-Dunne, Munck, & Ollendick, 2012; Feldman, 2009). Similarities across definitions put forth in the literature suggest that emotion regulation can be defined as a blend of the

physiological, behavioral, and cognitive processes responsible for monitoring, evaluating, and modulating the experience and expression of emotions in order to accomplish personal goals (Calkins & Hill, 2007; Cole et al., 2004; Eisenberg & Spinrad, 2004; Gross & Thompson, 2007). The richness in this definition forecasts the challenges in operationalizing and measuring emotion regulation processes. This complexity is rooted in the intricate nature of emotions themselves.

Emotions are ongoing processes that unfold over time. According to the modal model of emotion regulation (also known as the process-oriented approach), emotions are generated in a four-step process: a psychologically relevant situation takes place that is then attended to and appraised by the individual and that appraisal activates an inclination toward a particular course of action (Gross & Thompson, 2007). Emotion regulation can occur at different moments during the emotion-generation process. *Antecedent-focused* strategies are those adopted before emotions become fully activated, whereas *response-focused* strategies are implemented *once* emotions are activated (Gullone, Hughes, King, & Tonge, 2010). Emotions also involve changes in multiple systems (e.g., thoughts, physiology, and behavior). Accordingly, there are strategies that aim to regulate emotion by focusing on mental processes (i.e., cognitive or meta-cognitive strategies) and strategies that focus on behaviors (Parkinson & Totterdell, 1999). Further distinctions have been made between *emotion-focused* strategies that attempt to modify the emotion itself and *problem-focused* strategies that attempt to modify the circumstances leading to the emotion (Lazarus & Folkman, 1987). Similarly, there are strategies that engage versus those that disengage with the emotion-eliciting event. In sum, over 300 emotion regulation strategies have been identified in the literature (Augustine & Hemenover, 2009; Parkinson & Totterdell, 1999; Skinner, Edge, Altman, & Sherwood, 2003). A comprehensive review and critique of different classification systems for emotion regulation strategies can be found in Skinner et al. (2003).

RULER capitalizes on the diversity of emotion regulation strategies and explicitly teaches children and adults how and when to use various strategies to manage their emotions and accomplish personal goals. Before describing how RULER does this, in the next section we review what is known about the development of emotion regulation, to highlight the potential of and challenges in teaching emotion regulation to children of different ages.

13.2 Emotion Regulation Development from Infancy to Childhood

Given the large number of strategies that can be used to regulate emotions, it is worth asking why so many children fail to do so in adaptive ways. One reason is that emotion regulation is a process acquired and refined across development, which may account for differences in individuals' use of regulation strategies (Dodge & Garber, 1991). The processes underpinning emotion regulation begin to develop prenatally and continue evolving throughout life (Diamond & Aspinwall, 2003).

A general developmental sequence has been mapped out from infancy through early childhood (up to age 5). Infants start with basic regulatory behaviors that are innate and reflexive, such as sucking and hand-to-mouth movements (Cole et al., 2004). Those basic regulatory behaviors are supplemented with interactions with caregivers, who provide important regulatory functions during the first months of life (Campos, Frankel, & Camras, 2004). Over time, basic reflexes, and behaviors learned through repeated interactions with caregivers, become incorporated into the infants' voluntary repertoire of responses (Cole et al., 2004)

Most typically developing children acquire a broad repertoire of self-regulatory strategies from ages 2–5 as their brains develop, the prefrontal cortex in particular, and cognitive, motor, and language skills advance (Cole et al., 2004). Awareness of emotion regulation strategies likely emerges between ages 3 and 5, when children begin to understand the nature of the mind and the relation between internal states and external behavior (Cole, Dennis, Smith-Simon, & Cohen, 2009). By age 5, children are capable of deliberate and planful regulation strategies, including the ability to distract themselves and use problem-focused coping skills, such as taking action to solve a problem (Kalpidou, Power, Cherry, & Gottfried, 2004). There is some evidence that although preschoolers have emerging intuitions about the effectiveness of different regulation strategies, they still have difficulties discerning between effective and ineffective regulation strategies (Denham, 1998; e.g., Saarni, 1999). For example, when presented with puppet characters expressing different negative emotions (anger, sadness, fear), 3- and 4-year-olds were able to correctly identify strategies that would be more effective at decreasing negative emotions, such as fixing the situation in some way or distracting oneself, and those that would be less effective, such as ruminating or venting (Dennis & Kelemen, 2009). However, researchers also reported that children of the same age viewed rumination and aggression as effective regulation strategies (Denham, 1998; Saarni, 1999).

The ability to manage emotions by using cognitive strategies also increases with age (Altshuler & Ruble, 1989; see also Chaps. 11 and 12). Cognitive strategies involve awareness that goals, thoughts, and emotions are interdependent and that changing one's goals (e.g., deciding to want something else) and thoughts (e.g., deciding to think about something pleasant) can change one's emotions (Davis, Levine, Lench, & Quas, 2010). Given the abilities underlying these kinds of strategies, it is reasonable to expect that children would have more difficulty than adults in implementing such strategies to manage their emotions (see also Casey, Getz, & Galvan, 2008). There is, however, little clarity regarding the specific ages at which children begin to understand and are able to use cognitive strategies (Davis et al., 2010). Some studies have found that compared to older children, 5–7-year-olds are less likely to spontaneously suggest cognitive strategies and are more likely to rate behavioral strategies as more effective than cognitive strategies to regulate emotion (Brown, Covell, & Abramovitch, 1991; Pons, Harris, & de Rosnay, 2004). Other studies have found that by age 5 children can describe, understand, and use cognitive regulation strategies (Davis et al., 2010). The reasons underlying these divergent findings require further investigation, but suggest that preschoolers' ability to reason and use cognitive strategies may vary as a function of the situation or the task at

hand and the degree of exposure to these strategies at home or school. Overall, it is clear that by the time children enter school their regulatory skills are still undergoing important changes.

Children's emotion regulation skills continue to evolve during middle and late childhood. This positive trend may result from a range of processes: Children's neural systems continue to mature; they interact with parents and siblings, navigate new contexts (such as daycare and school), and engage in relationships with peers and adults outside of the home (Davis & Levine, 2013). However, relatively little is known about the development of emotion regulation skills during the elementary school years (Adrian, Zeman, & Veits, 2011; Bridges, Denham, & Ganiban, 2004; Campos et al., 2004). We know that developmental trajectories beyond early childhood may vary across at least five dimensions: (1) *emotion differentiation*, e.g., is sadness or anger or excitement being regulated; (2) *focus of regulation*, e.g., is it the situation that needs to be modified or is it the emotion; (3) *the component of the emotional system being regulated*, e.g., regulation efforts focused on physiological vs. behavioral responses; (4) *type of strategy*, e.g., does the individual use cognitive or behavioral strategies; and (5) *type of display rule*, e.g., knowing when and how to express emotions verbally or nonverbally.

Developmental research suggests different trajectories of skills across these dimensions. By way of illustration, consider the first dimension, emotion differentiation. Five-year-olds have a well-developed understanding of basic emotions such as sadness and fear, but only after 7 years of age come to understand more complex emotions such as shame and embarrassment (Widen & Russell, 2010). Similarly, for the type of strategy dimension, 9–10-year-olds, but not younger children, understand that some cognitive strategies, like dampening of expectations or being pessimistic, are useful for preventing only certain emotions, such as the emotion of disappointment (Guttentag & Ferrell, 2008). Also, regarding the type of display rule dimension, 10-year-olds have a basic understanding of how and when to express emotions verbally, but even 15-year-olds have difficulty mastering the rules governing the facial display of emotions (e.g., masking disappointment to avoid hurting someone's feelings; Gnepp & Hess, 1986). In general, it is expected that with maturation and experience, children expand their repertoire of regulatory strategies. By middle to late childhood, most children can generate a great number of approaches to deal with negative situations and to manage negative emotions; they can assess more accurately the degree of control over a situation, intentionally shift their attention when a situation cannot be modified, reappraise the meaning of the situation, analyze a situation from multiple angles, and problem-solve from different perspectives (Saarni, 1997).

Not all children who are theoretically capable of regulating the experience and expression of emotions develop and are able to implement effective strategies (e.g., Blandon, Calkins, Keane, & O'Brien, 2008). Children may lack opportunities to observe and learn effective strategies in the home and other contexts. They may be unable to retrieve strategies stored in memory when faced with an emotionally challenging situation or have difficulties choosing the right strategy to match the demands of the situation. Similar to the acquisition of other skills, it is possible that

explicit and sequenced training and continuous opportunities for practice may increase the likelihood of mastering and using effective regulation strategies across different situations.

School contexts are ripe with opportunities to nurture the development of children's emotion regulation skills. Classrooms, in particular, are settings where children acquire and reinforce adaptive or maladaptive skills, behaviors, and attitudes, by means of observing and interacting with their peers and teachers. RULER, and other evidence-based approaches to SEL, can make a difference in supporting children's emerging emotional intelligence (CASEL, 2013). By taking advantage of the instructional and social interactions already occurring in the classroom, RULER helps educators to maximize the opportunities for developing children's emotional intelligence by making SEL instruction fully integrated into the daily curriculum in an intentional, structured, and consistent way. We now turn to a more detailed description of RULER.

13.2.1 Teaching Emotional Intelligence with RULER

RULER is a universal approach to SEL designed to enhance the emotional intelligence of all members of the school community (children and adults), across all grade levels. The approach is premised on the idea that integrating emotional intelligence into the core academic curriculum and providing training and supports for all school members are necessary conditions for children to succeed socially, emotionally, and academically (Brackett et al., 2009). Guided by the ability model of emotional intelligence (Mayer & Salovey, 1997), RULER targets five interdependent skills that can be learned through experience and instruction, as opposed to being innate capacities or traits, and that are central for important developmental outcomes. The skills are recognizing, understanding, labeling, expressing, and regulating emotions (which form the acronym RULER).[1] We briefly describe each of these skills below, not only to provide a sense of the scope of RULER but also because those skills are building blocks for the development of emotion regulation. For more in-depth descriptions, see Brackett et al. (2011).

Recognizing emotions entails attending to and accurately interpreting relevant nonverbal cues in others and in the self and knowing how such cues relate to emotional states. Cues may include thoughts, physiological changes, facial expressions, tone of voice, and behavior (Ekman, 2003). Accurate recognition of emotional

[1] Mayer and Salovey's model identifies four emotional intelligence skills: perceiving, using, understanding, and managing emotions. RULER builds the same set of skills but uses different labels to facilitate comprehension and recall within an education context. For example, recognizing emotion in RULER is similar to perceiving emotion in Mayer and Salovey's model, with both referring to the accurate identification of emotions in the self and others. However, RULER distinguishes between "understanding" the causes and consequences of emotions and "labeling" emotions using a sophisticated vocabulary, whereas Mayer and Salovey combine these two areas into their understanding emotion branch.

experiences is fundamental for selecting appropriate courses of action (Denham et al., 2003). For example, a child who interprets a peer's anxiety as anger may respond aggressively instead of supportively.

Understanding emotions requires awareness about their causes and consequences and knowledge about how emotions combine, progress, and transition from one another. Children who understand emotions are better able to anticipate and recognize emotions that are likely to arise in a given situation (Stein & Levine, 1999). For example, a child who knows that falling behind with homework will make her feel overwhelmed is more likely to work harder to prevent that situation.

Labeling emotions refers to the availability and use of a rich emotion vocabulary to describe the full range of emotional experiences. Accurate labeling of emotions enhances children's ability to communicate about their feelings and build more harmonious social relationships (Denham, 1998). The ability to accurately recognize and label emotions has been associated with positive social behaviors and academic attainment over time (Izard et al., 2001).

Expressing emotions in socially appropriate ways involves an understanding of display rules or the implicit and explicit norms that dictate when and how to express emotions (Saarni, 1979). Children with a better understanding of display rules are rated as more socially competent by both teachers and peers (McDowell, O'Neil, & Parke, 2000).

Regulating emotion, as discussed previously, refers to the collection of strategies that individuals may use when attempting to manage their emotional states, in the service of personal goals (Gross & Thompson, 2007). Effective emotion regulation has been linked to a diversity of highly valued outcomes, including academic attainment, peer relationships, and mental health (Calkins, Gill, Johnson, & Smith, 1999; Eisenberg et al., 1997; Gumora & Arsenio, 2002; Werner & Gross, 2010).

In the following sections, we describe how RULER embeds the teaching of emotional intelligence, and emotion regulation skills in particular, throughout the academic curricula.

13.2.2 RULER's Approach for Teaching Emotion Regulation

RULER embeds the learning and practicing of emotional intelligence throughout the academic curricula and the ethos of a school. To accomplish this goal, it uses four "anchor" tools and the Feeling Words Curriculum, an advanced program designed to integrate into existing K-8 curricula (for a comprehensive description, see Brackett et al., 2011). To fully understand these methods, it is important to understand the theory of change underlying RULER.

RULER's theory of change, as illustrated in Fig. 13.1, is founded primarily on two psychological theories: the ability model of emotional intelligence (Mayer & Salovey, 1997; Salovey & Mayer, 1990) and ecological systems theory (Bronfenbrenner & Morris, 2006). The theory of change articulates that students learn and apply the skills of emotional intelligence through diverse developmentally appropriate activities such

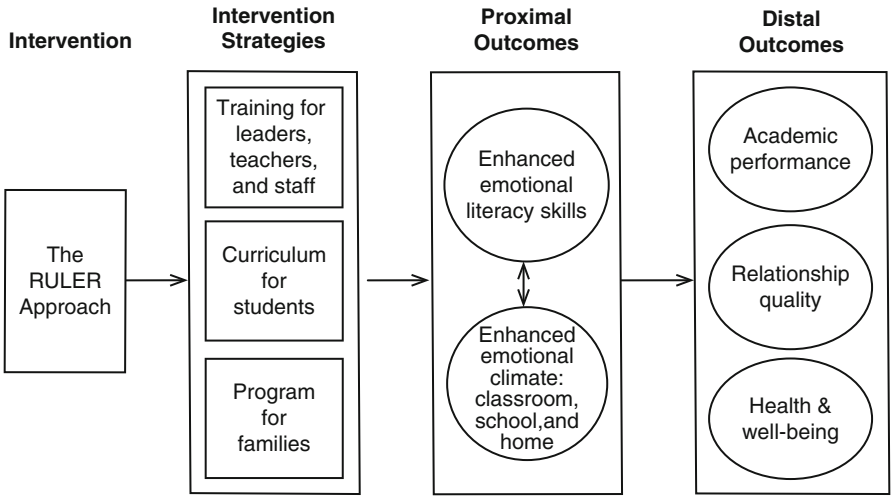

Fig. 13.1 The RULER approach theory of change for students

as self-reflection, classroom discussions, and creative writing (Brackett et al., 2011). In addition to nurturing individual teachers and students' skills, RULER also transforms school-wide academic and social interactions between and among children and adults. RULER contends that teachers, principals, and others adults serve as role models for children and that their skills also shape the social and instructional interactions at school. It is for this reason that school administrators and educators participate in their own professional development so they can develop their own skills and learn best practices for implementing and sustaining the program. More engaging and caring learning environments will arise when all members of the school community improve their emotional intelligence and when RULER is implemented with the goal of integrating into the curriculum and all school routines.

RULER provides continuous opportunities for all school members, adults and children alike, to learn and apply the skills of emotional intelligence (Brackett et al., 2009). By shifting the attitudes and skills of all school members, RULER aims to transform the quality of children's learning and social interactions. Those whole school changes, in turn, are expected to enhance academic performance, social relationships, and overall mental and physical health (Brackett & Rivers, 2014). A recent randomized control trial in 64 classrooms showed that RULER led to significant improvements in the quality of classrooms' social and pedagogical interactions (Hagelskamp, Brackett, Rivers, & Salovey, 2013; Rivers, Brackett, Reyes, Elbertson, & Salovey, 2013). These findings add to the growing body of evidence showing that programs like RULER can be effective approaches to improving children's academic and psychosocial outcomes (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011).

RULER's tools and pedagogical techniques contribute to developing in children and adults the skills of recognizing, understanding, labeling, expressing, and regulating

emotions. Here, we briefly describe each of the tools and techniques with a special emphasis on their role in teaching emotion regulation (for a comprehensive description, see Brackett, Caruso, & Stern, 2013).

The Charter *The Charter* is an agreement developed jointly by students and educators about how they want to feel in the classroom and what everyone can do to nurture those feelings (Brackett et al., 2013). In creating the classroom Charter, all classroom members follow four general steps. First, they reflect about their own feelings and the feelings of others. This step promotes self- and social-awareness, both of which are necessary for emotion regulation (Hoeksma, Oosterlaan, & Schipper, 2004). Second, they brainstorm about concrete and observable behaviors associated with each feeling. The process of linking behaviors and feelings builds on and promotes students' understanding of the causes and consequences of emotions. Third, everyone shares behavioral and cognitive strategies that can be used to handle uncomfortable feelings and behaviors. Sharing strategies as a group expands the repertoire of responses available to each individual student, facilitating the use of adaptive responses. In the fourth and final step, the jointly created vision is signed by all members of the classroom and posted for everyone to see and reference. Creating and regularly using the classroom Charter is expected to foster a positive emotional climate that everyone is accountable for. Classroom environments that are emotionally safe and supportive allow opportunities for students to observe and practice adaptive emotion regulation strategies, thus contributing to students' social and emotional growth (Jennings & Greenberg, 2009).

The Mood Meter The *Mood Meter* is a visual tool that allows educators and students to plot their emotional states according to two components of emotions: pleasantness and energy (see Fig. 13.2). Educators and students choose a number ranging from negative five (-5) to positive five (+5) to indicate their levels of subjective pleasantness (very unpleasant to very pleasant) and energy (no energy at all to a lot of energy). The point where pleasantness and energy cross corresponds to one of four quadrants. The *red* quadrant (upper left) is characterized by unpleasant/high-energy emotions, such as anger and anxiety. The *blue* quadrant (lower left) represents unpleasant/low-energy emotions, such as boredom and desolation. The *green* quadrant (lower right) is for pleasant/low-energy emotions, such as peacefulness and serenity. And the yellow quadrant (upper right) is characterized by pleasant/high-energy emotions, such as joy and elation. Once the plot is determined, educators and students choose an emotion word to describe their emotional state. Plotting emotions in the Mood Meter helps build emotional awareness about the self and others, and finding words to describe emotions promotes growth in labeling emotions with a rich vocabulary. Students and educators also can use the Mood Meter to consider their goals and whether a given emotional state is conducive to achieving those goals. For example, emotions in the yellow quadrant facilitate the generation of creative ideas and cooperative activities, whereas emotions in the green quadrant are useful for reflecting and reaching consensus. As educators and students reflect about the role that emotions play in everyday activities, they deepen their understanding about the consequences of emotions.

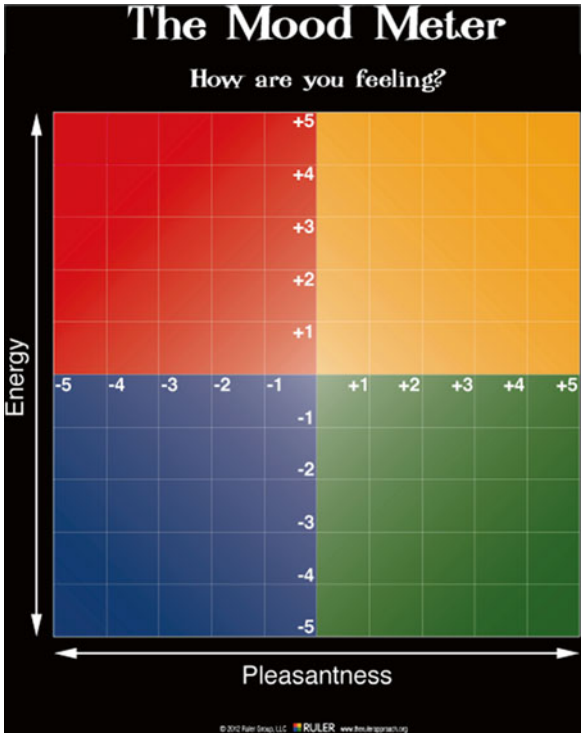

Fig. 13.2 The Mood Meter

Within the RULER framework, students and educators also learn about concrete emotion regulation goals and strategies. To facilitate learning of different categories of regulatory goals, students are introduced to the acronym PRIME: Students learn that emotion goals can include **p**reventing (i.e., not feeling something), **r**educing (i.e., feeling less), **i**nitiating (i.e., feeling something different), **m**aintaining (i.e., feeling the same), or **e**nhancing (i.e., feeling more) emotions. The strategies are informed by the typologies found in the literature (Gross & Thompson, 2007; Parkinson & Totterdell, 1999) and are categorized as thought vs. action strategies, which makes them more intuitive for teachers and students. Such strategies include breathing, distraction, positive reappraisal, and acceptance. With the Mood Meter as a guide, students and teachers can discuss if and how they want to shift their feeling state on the Mood Meter. Teachers also help students to evaluate ineffective strategies such as acting out, rumination, and procrastination to regulate emotions. Teachers use developmentally appropriate techniques for introducing strategies so that even young children can begin to recognize them and work to eliminate them from their repertoire.

Meta-Moment The *Meta-Moment* teaches children and adults how to avoid reacting to emotional triggers (e.g., an insult, negative feedback, bad news) using ineffective strategies (e.g., screaming, acting aggressively, withdrawing), but instead using effective strategies (e.g., positive self-talk, reappraisal). The technique involves a

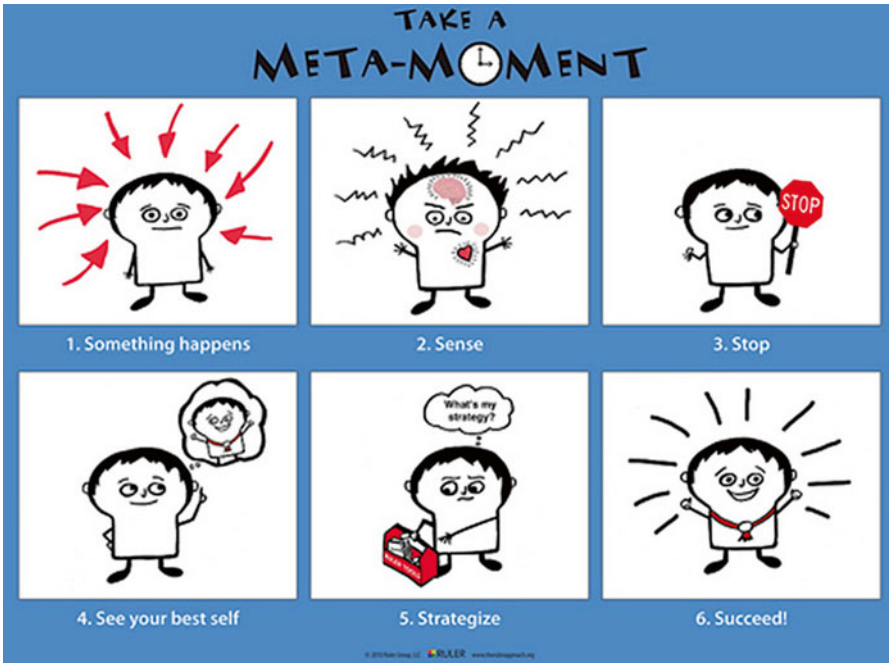

Fig. 13.3 The six steps in the Meta-Moment

process for prolonging and managing the time between an emotional trigger and a response. There are six steps to the Meta-Moment:

1. Realizing an emotional trigger has happened
2. Recognizing that an emotional response is happening – noticing changes in one's cognitive, physiological, and behavioral activity
3. Taking deep breaths in order to pause and avoid an impulsive response to the trigger
4. Seeing one's "best self" to shift attention away from the "object" or person
5. Choosing and implementing an effective emotion regulation strategy
6. Succeeding in applying an effective strategy by behaving in a way that is aligned with one's best self

Figure 13.3 includes a graphical depiction of the Meta-Moment process that teachers use in their classrooms. When teaching the Meta-Moment, the teacher takes students through the six steps by asking them to think about situations that trigger strong emotions. When first learning the process, students develop awareness about their emotions by thinking about the causes of emotion (Step 1) and by learning to identify subtle changes (e.g., thoughts, bodily reactions) in themselves and others (Step 2). Students learn how to breathe and pause in order to calm their minds and bodies by activating their parasympathetic nervous system (Step 3). They learn that pausing is an opportunity to choose a response that is best aligned

with their personal goals. Teachers help students to describe their "best self" in words (for older students) or in drawings (for younger ones) (Step 4). The best self includes students' self-image, as well as how they want to be seen by others, and the outcomes they value and want to achieve. Through collaborative work groups and discussions, teachers help students identify concrete regulation strategies to effectively manage their emotions and behave in ways that are congruent with their best selves (Step 5). The last step of the process involves the student succeeding in behaving in a way that aligns with his or her best self (Step 6).

By introducing the process and reinforcing it through regular practice and modeling, students learn that when an emotion is triggered, they can take a deep breath and visualize their best self. These steps enable them to generate responses that help them achieve personal goals and avoid responses that they may later regret. In other words, students think about what it looks like to behave in a way that is consistent with their best self. This positive visualization serves as an incentive and reinforcement for choosing an effective strategy and positive behavioral response. Visualization of best possible selves has been associated with positive emotion regulation and well-being outcomes (Owens & Patterson, 2013; Sheldon & Lyumbomirsky, 2006). Once adults and children learn the process, they are encouraged to engage actively in taking Meta-Moments throughout the day, which results in adult modeling and both adult-to-child and child-to-child coaching on taking Meta-Moments. Even the use of the term "Meta-Moment" provides a shared language and norms around emotion regulation that all members of the school community use in daily interactions.

Blueprint The *Blueprint* is a problem-solving tool that can be applied to solve past, present, and future interpersonal conflicts and analyze situations involving oneself or someone else, including characters in history and literature. The Blueprint guides educators and/or students through the process of identifying and describing: what emotions are experienced by each person, the causes of each person's feelings, how each person expresses and manages those feelings, and what each person could have done, or can do presently or in the future, to handle their emotions more effectively. The Blueprint employs all emotion skills, is useful in fostering empathy and perspective taking, and strengthens regulation skills by encouraging educators and students to critically assess the results of ineffective regulation strategies and visualize more adaptive strategies.

Feeling Words Curriculum RULER's *Feeling Words Curriculum* (FWC) integrates into the standard classroom curriculum from preschool through eighth grade. It consists of five steps that teachers use to embed between 15 and 20 new *feeling words* into their curriculum throughout the school year. The feeling words that comprise the FWC describe a full range of emotional experiences from joy to grief to frustration to contentment to excitement. The curriculum offers a set of words that span development and build from the previous year's words such that kindergartners learn *angry*, third graders learn the more advanced *annoyed*, and fifth graders learn *irate*.

The first four steps encourage students to:

1. Connect their own personal experiences with the new vocabulary (i.e., the personal association)
2. Make an explicit link between the new word and the academic content (i.e., the academic link)
3. Introduce the word to adults in their homes (i.e., the school-home partnership)
4. Explore the word through creative expression, such as the visual and performing arts (i.e., the creative connection)

Collectively, these four steps expand children's feelings vocabulary – which has been shown to facilitate emotion regulation efforts (Lieberman et al., 2007) – and help students become more attuned to their own emotions and those of others, strengthen their knowledge about how emotions look and feel, their causes and consequences, and about how emotions can be more appropriately expressed (Brackett et al., 2010). These skills help children to be better equipped to employ situation selection strategies that require knowledge and understanding of the causes and consequences of emotions.

The four steps also serve as stepping-stones for the fifth step, the (5) *strategy session*, which explicitly reinforces students' knowledge of emotion regulation strategies. The strategy session is based on the idea that children need opportunities to both reflect on and problem-solve about strategies – both effective and ineffective – that can be used to regulate emotions. This step is designed to expose students to multiple perspectives and allows them to learn both cognitive strategies (e.g., focusing on positive aspects of the situation, acceptance, visualization) and behavioral strategies (e.g., breathing, relaxing, engaging in physical activity, finding support from others) for managing emotions in themselves and others. Teachers present students with scenarios that describe an emotion in need of regulation. The scenario may be hypothetical, pulled from a piece of literature the students are studying, from current events, or from their personal experiences (such as something that happened in the classroom or the playground). The teacher facilitates a collaborative discussion and work session whereby students identify the emotion goals in the scenario (e.g., how does the person want to feel) and emotion regulation strategies (e.g., what can the person think about or do to feel that way). Students then brainstorm about strategies that can be used to achieve the selected goal. After students' ideas are recorded, a discussion follows where students analyze why certain strategies may be more or less effective in a given situation. Students in higher grades reflect on which strategies would work best for them and why.

The sequenced activities in the FWC provide opportunities for repeated exposure, reflection, and practice of regulation strategies across a wide range of feelings (from grief to satisfaction to alienation to fury to elation). The activities are developmentally appropriate and designed to meet the learning needs and capacity at different grade levels. For example, elementary school students are presented with strategies to experience "more, less, or the same" amount of an emotion, upper elementary school students learn how to differentiate thought and action strategies, and middle school students consider the effectiveness of strategies in both the short and

long term. The activities are designed to expand students' repertoire, prepare them to deal with diverse situations, enable them to select strategies aligned with their social and academic goals, and increase accessibility to effective instead of ineffective regulation strategies. Together, all the components of RULER equip students with the tools and skills that they require to effectively regulate their emotions in a variety of scenarios.

13.2.3 Measuring Emotion Regulation Skills

Assessing emotional intelligence across development and in response to SEL programming will provide important information for educators, policy-makers, and researchers alike. Most importantly, tracking students' emotional skill development can help teachers tailor teaching to meet specific needs and help researchers modify SEL programming to be more effective at promoting positive youth development. In the following section, we discuss how measuring emotion regulation skills, in particular, can inform educators' practices, as well as the effectiveness of programs like RULER in fostering children's positive development.

The measurement of emotion regulation for research purposes has made important strides in recent years (Adrian et al., 2011). Observational paradigms have been developed that closely resemble social situations, such as peer rejection and provocation, which trigger strong emotional responses in children and adolescents (e.g., Reijntjes, Stegge, Terwogt, Kamphuis, & Telch, 2006; Underwood, Hurley, Johanson, & Mosley, 1999). These innovative paradigms permit assessing children's emotional responses and regulation strategies in real time. Also, technological advances, such as those involved in the measurement of rhythmicity of respiratory sinus arrhythmia (RSA), are enabling a more precise understanding of the physiological and biological processes that underlie emotional responses and strategies (Porges, 2007). Despite these significant advances, further work is needed to develop accessible measures that can inform the work of educators and SEL program developers (Zeman, Klimes Dougan, Cassano, & Adrian, 2007). Here, we discuss ways in which formative assessments of emotion regulation can be geared toward the enhancement of student outcomes.

Formative assessments have gained popularity as a way to fine-tune teaching and learning to match students' needs (Black & Wiliam, 1998). In contrast to other kinds of assessments designed to grade, certify, or rank students, the primary objective of formative assessments is to help students and teachers make adjustments to instructional strategies based on what students know and are yet to learn, relative to specific learning goals. In an extensive review of the literature, teachers' use of formative assessments was found to predict significant gains in academic outcomes, particularly for low-achieving students (Black & Wiliam, 1998). Yet, to the best of our knowledge, the potential of formative assessments to enhance the teaching and learning of social and emotional skills awaits to be explored. Formative assessments can be useful in teaching these skills, as they can aid teachers in identifying students'

strengths and weaknesses, planning and focusing future lessons and activities, providing more specific and grounded feedback, and fostering students' ownership over their learning process (Andrade & Cizek, 2010). In this section we provide a few concrete examples of the ways in which formative assessments can be used to strengthen the teaching of emotion regulation and the implementation of RULER and other SEL programs.

RULER activities are designed to seamlessly include formative assessments. For example, in teaching the "strategize" step of the Meta-Moment, teachers introduce evidence-based emotion regulation strategies and ask students to think critically about how they can use each strategy to deal more effectively with their "triggers." By way of example, after introducing positive self-talk, third grade teachers ask their students to brainstorm about specific things they could say to themselves to feel less annoyed, nervous, or sad using both hypothetical and personal scenarios. Students are prompted to justify their responses, giving them opportunities to reevaluate the usefulness of ineffective strategies (Brackett et al., 2013). At the end of the lesson, teachers can assess student learning by presenting new scenarios to students and then asking them to select from a list the most effective strategy to regulate the emotion. Choices can include both effective and ineffective strategies. Teachers then have data on student understanding to guide what lesson components need to be revisited with the whole class or with select students who require additional work to grasp the concepts. For example, some children might be better able to identify effective strategies for regulating anger as opposed to sadness.

An important aspect of formative assessments is to empower students to take ownership over their own learning process (Andrade & Cizek, 2010). RULER encourages students (and teachers) to practice Meta-Moments every day. As part of the program, teachers ask students to choose a specific situation that they find difficult to manage and one that also consistently activates an unpleasant emotion (e.g., being called names by a peer). In upper elementary classrooms (third through fifth grades), teachers ask students to fill out a worksheet that takes them through the six steps in the Meta-Moment (see Fig. 13.3). After students complete the worksheet, the teacher asks them to set a goal for enacting the Meta-Moment in the same or a similar situation. At the end of the week, teachers ask students to write about their successes and challenges. Teachers collect students' responses and use them to provide feedback about alternative strategies that students can use to succeed or to reinforce the use of effective strategies (Brackett et al., 2013). In order to further promote empowerment, teachers may also ask students to identify new ideas about steps they can take to succeed the next time they face similar, emotionally laden situations.

Further, existing validated measures of emotion regulation can be adapted for the purposes of formative assessment. One such measure is the 12-item Children's Sadness Management Scale (CSMS) designed to assess children's expression and regulation of sadness (e.g., "I hold my sadness in," "I whine/fuss about what's making me sad," "I stay calm and don't let sad things get to me," "When I'm sad, I do something totally different until I calm down") (Zeman, Shipman, & Penza-Clyve, 2001). Normed with 9–12-year-olds, teachers can administer the self-report CSMS

to all students at the beginning of the year, to obtain a baseline for the class and for individual children. The baseline, along with repeated measures administered during the course of the year, can be used to track children's growth and to calibrate the focus of Meta-Moment activities to match the group's initial strengths and weaknesses. The same data could be used by the developers of RULER to identify groups of children who are likely to benefit the most from the program as is, as well as groups who may require additional supports. Using this information, program activities can be refined to better meet the needs of diverse groups of children. Further, program developers can combine data on students' change trajectories with implementation fidelity data (e.g., number of lessons taught by teachers in a school), to determine the optimal level of implementation that is typically conducive to greater changes in children's use of effective emotion regulation strategies. Other measures that could be integrated with a formative assessment protocol for emotion regulation are the Children's Worry and Anger Management Scales (CWMS, CAMS), which have similar features as the CSMS and have been validated for children of the same ages (Zeman, Cassano, Suveg, & Shipman, 2010).

Systematic data collection of children's skills for using and regulating their emotions as part of routine classroom activities can improve our understanding of the developmental trajectories of these skills. Knowledge about what children typically do, as well as what they could do if provided with adequate supports, can serve to establish developmental benchmarks and set more ambitious yet attainable goals for educational programs. In the absence of detailed information about typical development, it is difficult to contextualize the gains made in the presence of SEL programs. Information also can be gathered on the factors that lead to difficulties in the regulation of emotions, such as characteristics of the situation (e.g., a test, recess), the group (e.g., number of children with behavior problems in a classroom), and the individual (e.g., age, cultural background). This information will contribute to a deeper understanding of individual differences in acquiring and implementing effective regulation strategies and about how to support children with different learning styles and needs.

13.2.4 Conclusion

Emotional intelligence plays a significant role in all spheres of life. Yet, the development of those skills is often left to chance, and few of us encounter structured, systematic, and deliberate opportunities for their improvement. For that reason, many children and adults lack the skills necessary to lead happier and more productive lives.

The number of SEL programs that focus, at least in part, on the development of children's skills has increased steadily over the last few decades. This growth is testament to a positive shift in the value that we, as a society, place on emotional intelligence. However, there is progress to be made in the use of data-driven approaches for the improvement of specific teaching practices and in the evaluation of SEL programs.

As illustrated in this chapter, SEL programs founded on theory and research have the potential to exert positive changes in children's lives. However, our review of the literature on emotion regulation suggests that more research is needed to better understand how children's emotional intelligence continues to evolve throughout elementary school, the barriers that some children face in developing positive skills, and what children can achieve with the support of high-quality SEL programs. Expanding and refining what we know about the development of emotional intelligence will permit the design of more effective and efficient SEL programs.

Research on the use of formative assessments to boost academic learning also points to promising applications for the teaching and learning of emotional intelligence. As discussed in the last section of the chapter, collecting data on children's developing emotional intelligence in schools can serve multiple purposes. Teachers who are aware of their students' strengths and limitations can make better decisions about lesson plans and ascertain the need for additional or more focused supports. The same data can be used by program developers to determine whether the program is more or less effective for certain students, identify optimal and suboptimal conditions for program implementation, and make adjustments to increase the likelihood of successful implementation and outcomes. The use of formative assessments to aid in fostering children's emotional intelligence is an area that awaits development.

In summary, efforts to expand the knowledge base about the development of emotional intelligence and to develop formative assessments to improve the practice, implementation, and effectiveness of SEL programs can help advance the field of SEL and support all children in reaching their full potential.

References

Adrian, M., Zeman, J., & Veits, G. (2011). Methodological implications of the affect revolution: A 35-year review of emotion regulation assessment in children. *Journal of Experimental Child Psychology, 110*, 171–197.

Altshuler, J., & Ruble, D. (1989). Developmental changes in children's awareness of strategies for coping with uncontrollable stress. *Child Development, 60*(6), 1337–1349.

Andrade, H., & Cizek, G. J. (2010). *Handbook of formative assessment*. New York: Routledge.

Augustine, A. A., & Hemenover, S. H. (2009). On the relative effectiveness of affect regulation strategies: A meta-analysis. *Cognition and Emotion, 23*(6), 1181–1220.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7–74.

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57*, 111–127.

Blair, K. A., Denham, S. A., Kochanoff, A., & Whipple, B. (2004). Playing it cool: Temperament, emotion regulation, and social behavior in preschoolers. *Journal of School Psychology, 42*, 419–443.

Blandon, A. Y., Calkins, S. D., Keane, S. P., & O'Brien, M. (2008). Individual differences in trajectories of emotion regulation processes: The effects of maternal depressive symptomatology and children's physiological regulation. *Developmental Psychology, 44*(4), 1110–1123.

Brackett, M. A., Caruso, D. R., & Stern, R. S. (2013). *The anchors of emotional intelligence, K-8. The RULER approach to social and emotional learning*. New Haven, CT: Emotionally Intelligent Schools, LLC.

Brackett, M. A., Kremenitzer, J. P., Maurer, M., Rivers, S. E., Elberston, N. A., & Carpenter, M. D. (Eds.). (2011). *Creating emotional literate classrooms: An introduction to The RULER approach to social and emotional learning*. Port Chester, NY: National Professional Resources, Inc.

Brackett, M. A., Patti, J., Stern, R., Rivers, S. E., Elbertson, N. A., Chisholm, C., et al. (2009). A sustainable, skill-based approach to building emotionally literate schools. In D. Thompson, M. Hughes, & J. Terrell (Eds.), *The handbook of developing emotional and social intelligence: Best practices, case studies, & tools* (pp. 329–358). New York: Pfeiffer.

Brackett, M. A., Rivers, S. E., Holzer, A. A., Shapses, S., Elbertson, N. A., & Maurer, M. (2010). *The RULER approach: Feeling words curriculum (Grades K to 2)*. New Haven, CT: Ruler Group.

Bradley, R. T., McCraty, R., Atkinson, M., Tomasino, D., Daugherty, A., & Arguelles, L. (2010). Emotion self-regulation, psychophysiological coherence, and test anxiety: Results from an experiment using electrophysiological measures. *Applied Psychophysiology and Biofeedback, 35*(4), 261–283.

Bridges, L. J., Denham, S. A., & Ganiban, J. M. (2004). Definitional issues in emotion regulation research. *Child Development, 75*(2), 340–345.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology* (Theoretical models of human development 6th ed., Vol. 1, pp. 793–828). Hoboken, NJ: Wiley.

Brown, K., Covell, K., & Abramovitch, R. (1991). Time course and control of emotion. Age differences in understanding and recognition. *Merrill-Palmer Quarterly, 37*(2), 273–287.

Calkins, S. D., Gill, K. L., Johnson, M. C., & Smith, C. L. (1999). Emotional reactivity and emotional regulation strategies as predictors of social behavior with peers during toddlerhood. *Social Development, 8*(3), 310–334.

Calkins, S. D., & Hill, A. (2007). Caregiver influences on emerging emotion regulation: Biological and environmental transactions in early development. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 229–248). New York: Guilford.

Calkins, S. D., & Keane, S. P. (2004). Cardiac vagal regulation across the preschool period: Stability, continuity, and implications for childhood adjustment. *Developmental Psychobiology, 45*(3), 101–112.

Campos, J. J., Frankel, C. B., & Camras, L. (2004). On the nature of emotion regulation. *Child Development, 75*(2), 377–394.

Casey, B. J., Getz, S., & Galvan, A. (2008). The adolescent brain. *Developmental Review, 28*(1), 62–77.

Cole, P. M., Dennis, T. A., Smith-Simon, K. E., & Cohen, L. H. (2009). Preschoolers' emotion regulation strategy understanding: Relations with emotion socialization and child self-regulation. *Social Development, 18*(2), 324–352.

Cole, P. M., Martin, S. E., & Dennis, T. A. (2004). Emotion regulation as a scientific construct: Methodological challenges and directions for child development research. *Child Development, 75*(2), 317–333.

Collaborative for Academic, Social, and Emotional Learning. (2012). 2013 CASEL guide: Effective social and emotional learning programs—Preschool and elementary school edition. Chicago, IL: Author.

Davis, E. L., & Levine, L. J. (2013). Emotion regulation strategies that promote learning: Reappraisal enhances children's memory for educational information. *Child Development, 84*(1), 361–374.

Davis, E. L., Levine, L. J., Lench, H. C., & Quas, J. A. (2010). Metacognitive emotion regulation: Children's awareness that changing thoughts and goals can alleviate negative emotions. *Emotion, 10*(4), 498.

Denham, S. A. (1998). *Emotional development in young children*. New York: Guilford.

Denham, S. A., Blair, K. A., DeMulder, E., Levitas, J., Sawyer, K., Auerbach-Major, S., et al. (2003). Preschool emotional competence: Pathway to social competence. *Child Development, 74*, 238–256.

Dennis, T., & Kelemen, D. A. (2009). Preschool children's views on emotion regulation: Functional associations and implications for social-emotional adjustment. *Journal of Behavioral Development, 33*(3), 243–252.

Diamond, L. M., & Aspinwall, L. G. (2003). Emotion regulation across the life span: An integrative perspective emphasizing self-regulation, positive affect, and dyadic processes. *Motivation and Emotion, 27*(2), 125–156.

Dodge, K. A., & Garber, J. (1991). Domains of emotion regulation. In J. Garber & K. A. Dodge (Eds.), *The development of emotion regulation and dysregulation*. Cambridge, MA: Cambridge University Press.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*(1), 405–432.

Eggum, N. D., Eisenberg, N., Kao, K., Spinrad, T. L., Bolnick, R., Hofer, C., et al. (2011). Emotion understanding, theory of mind, and prosocial orientation: Relations over time in early childhood. *Journal of Positive Psychology, 6*(1), 4–16.

Eisenberg, N., Gurhrie, I. K., Fabes, R. A., Reiser, M., Murphy, B. C., Holgren, R., et al. (1997). The relations of regulation and emotionality to resiliency and competent social functioning in elementary school children. *Child Development, 68*(2), 295–311.

Eisenberg, N., & Spinrad, T. L. (2004). Emotion-related regulation: Sharpening the definition. *Child Development, 75*(2), 334–339.

Ekman, P. (2003). *Emotions revealed*. New York: Henry Holt and Company.

Esbjørn, B. H., Bender, P. K., Reinholdt-Dunne, M. L., Munck, L. A., & Ollendick, T. H. (2012). The development of anxiety disorders: Considering the contributions of attachment and emotion regulation. *Clinical Child and Family Psychology Review, 15*(2), 129–143.

Fabes, R. A., Eisenberg, N., Jones, S., Smith, M., Guthrie, I., Poulin, R., et al. (1999). Regulation, emotionality, and preschoolers' socially competent peer interactions. *Child Development, 70*(2), 432–442.

Feldman, R. (2009). The development of regulatory functions from birth to 5 years: Insights from premature infants. *Child Development, 80*(2), 544–561.

Fischer, K. W., Shaver, P. R., & Carnochan, P. (1990). How emotions develop and how they organise development. *Cognition and Emotion, 4*(2), 81–127.

Gnepp, J., & Hess, D. L. R. (1986). Children's understanding of verbal and facial display rules. *Developmental Psychology, 22*(1), 103–108.

Graziano, P. A., Reavis, R. D., Keane, S. P., & Calkins, S. D. (2007). The role of emotion regulation in children's early academic success. *Journal of School Psychology, 45*(1), 3–19.

Gross, J. J., & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations. In J.J. Gross (Ed.), *Handbook of Emotion Regulation* (pp. 3–24). New York: Guilford Press.

Gullone, E., Hughes, E. K., King, N. J., & Tonge, B. (2010). The normative development of emotion regulation strategy use in children and adolescents: A 2-year follow-up study. *Journal of Child Psychology and Psychiatry, 51*(5), 567–574.

Gumora, G., & Arsenio, W. F. (2002). Emotionality, emotion regulation, and school performance in middle school children. *Journal of School Psychology, 40*(5), 395–413.

Guttentag, R., & Ferrell, J. (2008). Children's understanding of anticipatory regret and disappointment. *Cognition and Emotion, 22*(5), 815–832.

Hagelskamp, C., Brackett, M. A., Rivers, S. E., & Salovey, P. (2013). Improving classroom quality with the RULER approach to social and emotional learning: Proximal and distal outcomes. *American Journal of Community Psychology, 51*(3-4), 530–543.

Hoeksma, J. B., Oosterlaan, J., & Schipper, E. M. (2004). Emotion regulation and the dynamics of feelings: A conceptual and methodological framework. *Child Development, 75*(2), 354–360.

Izard, C. E., Fine, S., Schultz, D., Mostow, A., Ackerman, B., & Youngstrom, E. (2001). Emotion knowledge as a predictor of social behavior and academic competence in children at risk. *Psychological Science, 12*(1), 18–23.

Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of Educational Research, 79*(1), 491–525.

Jerome, E. M., Hamre, B. K., & Pianta, R. C. (2009). Teacher–child relationships from kindergarten to sixth grade: Early childhood predictors of teacher-perceived conflict and closeness. *Social Development, 18*(4), 915–945.

Kalpidou, M., Power, T., Cherry, K., & Gottfried, N. (2004). Regulation of emotion and behavior among 3- and 5-year-olds. *The Journal of General Psychology, 131*(2), 159–178.

Koles, B., O'Connor, E., & McCartney, K. (2009). Teacher-child relationships in prekindergarten: The influences of child and teacher characteristics. *Journal of Early Childhood Teacher Education, 30*(1), 3–21.

Lazarus, R. S., & Folkman, S. (1987). Transactional theory and research on emotions and coping. *European Journal of Personality, 1*(3), 141–169.

Lieberman, M. D., Eisenberger, N. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H., & Way, B. M. (2007). Putting feelings into words: Affect labeling disrupts amygdala activity in response to affective stimuli. *Psychological Science, 18*(5), 421–428.

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. J. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications* (pp. 3–34). New York: Basic Books.

McDowell, D. J., O'Neil, R., & Parke, R. D. (2000). Display rule application in a disappointing situation and children's emotional reactivity: Relations with social competence. *Merrill-Palmer Quarterly, 46*(2), 306–324.

Owens, R. L., & Patterson, M. M. (2013). Positive psychological interventions for children: A comparison of gratitude and best possible selves approaches. *The Journal of Genetic Psychology, 174*(4), 403–428.

Parkinson, B., & Totterdell, P. (1999). Classifying affect-regulation strategies. *Cognition and Emotion, 13*(3), 277–303.

Pons, F., Harris, P. L., & de Rosnay, M. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology, 1*(2), 127–152.

Porges, S. W. (2007). The polyvagal perspective. *Biological Psychology, 74*(2), 116–143.

Raver, C. C. (2002). Emotions matter: Making the case for the role of young children's emotional development for early school readiness. *Social Policy Report, 16*, 3–18.

Reijntjes, A., Stegge, H., Terwogt, M., Kamphuis, J., & Telch, M. (2006). Children's coping with in vivo peer rejection: An experimental investigation. *Journal of Abnormal Child Psychology, 34*(6), 873–885.

Rimm-Kaufman, S. E., Pianta, R. C., & Cox, M. J. (2000). Teachers' judgments of problems in the transition to kindergarten. *Early Childhood Research Quarterly, 15*(2), 147–166.

Brackett, M. A. & Rivers, S. E. (2014). Transforming students' lives with social and emotional learning. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 368–388). New York: Taylor and Francis.

Rivers, S. E., Brackett, M. A., Reyes, M. R., Elbertson, N. A., & Salovey, P. (2013). Improving the social and emotional climate of classrooms: A clustered randomized controlled trial testing the RULER approach. *Prevention Science, 14*(1), 77–87.

Rosen, P. J., Milich, R., & Harris, M. J. (2012). Dysregulated negative emotional reactivity as a predictor of chronic peer victimization in childhood. *Aggressive Behavior, 38*(5), 414–427.

Saarni, C. (1979). Children's understanding of display rules for expressive behavior. *Developmental Psychology, 15*(4), 424.

Saarni, C. (1997). Emotional competence and self-regulation in childhood. In P. Salovey & D. J. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications*. New York: Basic Books.

Saarni, C. (1999). *The development of emotional competence*. New York: Guilford.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality, 9*(3), 185–211.

Sheldon, K. M., & Lyumbomirsky, S. (2006). How to increase and sustain positive emotion: The effects of expressing gratitude and visualizing best possible selves. *The Journal of Positive Psychology, 1*(2), 73–82.

Sherman, R. F. (2011). *Social and emotional learning action network white paper*. Paper presented at the Clinton Global Initiative annual meeting.

Shields, A. M., Cicchetti, D., & Dyan, R. M. (1994). The development of emotional and behavioral self-regulation and social competence among maltreated school-age children. *Development and Psychopathology, 6*, 57–75.

Skinner, E. A., Edge, K., Altman, J., & Sherwood, H. (2003). Searching for the structure of coping: A review and critique of category systems for classifying ways of coping. *Psychological Bulletin, 129*(2), 216.

Spilt, J. L., Hughes, J. N., Wu, J.-Y., & Kwok, O.-M. (2012). Dynamics of teacher-student relationships: Stability and change across elementary school and the influence on children's academic success. *Child Development, 83*(4), 1180–1195.

Stein, N. L., & Levine, L. J. (1999). The early emergence of emotional understanding and appraisal: Implications for theories of development. In T. Dalgleish & M. J. Power (Eds.), *Handbook of cognition and emotion* (pp. 383–408). New York: Wiley.

Sutton, R. E., & Wheatley, K. F. (2003). Teachers' emotions and teaching: A review of the literature and directions for future research. *Educational Psychology Review, 15*(4), 327–358.

Thompson, R. A., & Meyer, S. (2007). Socialization of emotion regulation in the family. In J.J. Gross (Ed.), *Handbook of emotion regulation* (pp. 249–268). New York: Guilford.

Trentacosta, C., & Izard, C. E. (2007). Kindergarten children's emotion competence as predictor of their academic competence in first grade. *Emotion, 7*, 77–88.

Underwood, M. K., Hurley, J. C., Johanson, C. A., & Mosley, J. E. (1999). An experimental, observational investigation of children's responses to peer provocation: Developmental and gender differences in middle childhood. *Child Development, 70*(6), 1428–1446.

Werner, K., & Gross, J. J. (2010). Emotion regulation and psychopathology: A conceptual framework. In A. M. Kring & D. M. Sloan (Eds.), *Emotion regulation and psychopathology: A transdiagnostic approach to etiology and treatment* (pp. 13–37). New York: Guilford.

Widen, S. C., & Russell, J. A. (2010). Children's scripts for social emotions: Causes and consequences are more central than are facial expressions. *British Journal of Developmental Psychology, 28*, 565–581.

Zeman, J., Cassano, M., Suveg, C., & Shipman, K. (2010). Initial validation of the children's worry management scale. *Journal of Child and Family Studies, 19*(4), 381–392.

Zeman, J., Klimes Dougan, B., Cassano, M., & Adrian, M. (2007). Measurement issues in emotion research with children and adolescents. *Clinical Psychology: Science and Practice, 14*(4), 377–401.

Zeman, J., Shipman, K., & Penza-Clyve, S. (2001). Development and initial validation of the children's sadness management scale. *Journal of Nonverbal Behavior, 25*(3), 187–205.

Chapter 14
Psychosocial Skills in Large-Scale Assessments: Trends, Challenges, and Policy Implications

Jonas P. Bertling, Francesca Borgonovi, and Debby E. Almonte

14.1 Introduction

Educational large-scale assessments vary with respect to the focus of the targeted constructs, the population under study, and the specific framework adopted (Naemi et al., 2013). Yet, these assessments share many common features. The primary goal of large-scale comparative studies has traditionally been to evaluate and monitor cross-national or cross-state differences in the extent to which students are able to solve a series of tests in specific academic subjects. International assessments such as the *Trends in International Mathematics and Science Study* (TIMSS), the *Progress in International Reading Literacy Study* (PIRLS), and the *Programme for International Student Assessment* (PISA) have measured students' performance in mathematics, reading, and science. Recent cycles of PISA also measured complex problem solving (PISA 2012), financial literacy (PISA 2012), and collaborative problem solving (PISA 2015). The *National Assessment of Educational Progress* (NAEP) in the United States has a broader scope covering additional domains beyond mathematics, reading, and science with assessments in areas such as social studies, the arts, writing, and technology and engineering literacy.

While the central focus of educational large-scale assessments remains to measure what students know and can do in key content domains, there is a growing recognition of other factors that play a vital role in making the most of the opportunities schools provide (Farrington et al., 2012; Kautz, Heckman, Diris, ter Weel, & Borghans, 2014; see Jencks, 1979, for one of the earlier seminal works raising

J.P. Bertling (✉) • D.E. Almonte
Educational Testing Service, Princeton, USA
e-mail: jbertling@ets.org; dalmonte@ets.org

F. Borgonovi
OECD, Paris, France
e-mail: Francesca.borgonovi@oecd.org

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_14

awareness for multiple determinants of success). In order to effectively meet the economic, political, and social demands for competencies, much more is in fact required of students and adults than cognitive proficiency (Levin, 2012). Educational researchers have argued, for example, that students need to be engaged, motivated, eager to learn new things, and feel that they can succeed (Christenson, Reschly, & Wylie, 2012). Numerous valuable attributes beyond achievement can be described that enable students to live fulfilled lives, meet challenges, and make the most of the available opportunities along the way (Schunk & Mullen, 2013; see also Chap. 1). Consequently, educational systems may be evaluated in terms of their capacity to develop various aspects of human potential, ranging from subject-specific achievement to socio-emotional, psychological, ethical, and behavioral aspects.

In recent decades, large-scale assessments of K-12 students have progressively shifted attention from a narrow definition of success based on test scores in specific academic subjects to a more holistic measurement of students' abilities and skills including psychosocial or noncognitive factors. In this chapter, we conceptualize and examine how these factors are measured and considered in the context of large-scale educational assessments of school-aged children. We highlight main trends in the changing role of noncognitive factors in this context and describe key challenges and policy implications, drawing on examples from PISA and NAEP.

14.2 Leading International and National Large-Scale Assessments

The *Programme for International Student Assessment* (PISA[1]) is a triennial study examining academic performance of samples of over 500,000 students in more than 60 countries worldwide, making it the largest international large-scale assessment. The first PISA cycle was introduced in 2000 and the most recent wave has been conducted in 2015. The main PISA assessment instrument consists of a 2-h test designed to measure students' ability to solve real-life problems in mathematics, reading, and science (Organisation for Economic Co-operation and Development [OECD], 2013b). Until the 2012 cycle, assessments were paper based. With the 2015 cycle the PISA transitioned to administering the assessment on computers. Additionally, the samples of students fill out a questionnaire aimed at collecting information on their family characteristics, household possessions, study habits, and dispositions toward school and specific academic subjects. Each PISA cycle focuses on one of the three assessed subject domains – reading, mathematics, and science[2] as a main domain while other subject areas are assessed as minor domains. The alternation of main and minor domains has two effects: first, the main domain receives a substantially larger portion of the allocated testing time (OECD, 2012, 2014b); and, second, the questionnaire that students are asked to complete at the end of the

[1] http://www.oecd.org/pisa/

[2] In 2000 and 2009, the main domain assessed was reading, in 2003 and 2012 it was mathematics, and in 2006 and 2015 it was science.

testing time explores in-depth students' engagement, attitudes, and dispositions toward the main subject domain. Therefore, in 2012 when the assessment was primarily focused on mathematics, the sampled students reported on their mathematics self-efficacy, self-concept, and anxiety, and in 2006, when science was the main assessment domain, the student questionnaire explored students' self-beliefs about science. PISA samples are representative of all students enrolled in schools at seventh grade or above that were between 15 years and 3 months and 16 years and 2 months old at the time of the assessment. PISA uses a two-stage sampling design, whereby schools represent the primary sampling unit, and within each selected school, a random sample of 35 students is selected (more details on the PISA sampling strategy can be found in OECD, 2010).

The *National Assessment of Educational Progress* (NAEP[3]) is the largest nationally representative and continuing assessment of what students in the United States know and can do in various subject areas. NAEP reading, mathematics, writing, and science assessments were first administered to samples of students on a national level in 1969. In 1990, NAEP assessments and questionnaires were first administered to samples at the state level allowing for state-by-state comparisons, similar to the cross-country comparisons carried out in PISA. Today, NAEP administers national, state, district, and long-term trend assessments in order to measure what students know in mathematics, reading, science, writing, the arts, civics, economics, geography, US history, and technology and engineering literacy. National or "main" NAEP assessments are administered to samples of fourth-, eighth-, and 12th-grade students. State and district NAEP assessments are administered to samples of fourth- and eighth-grade students. In some administration years, state NAEP assessments are also administered to 12th-grade samples. Each assessment takes approximately 60 min to complete. While NAEP is administered regularly for mathematics, reading, and science, assessments in other areas, such as the arts and social studies, are administered periodically. NAEP mathematics and reading assessments are administered every 2 years, while NAEP science assessments are administered every 4 years. Unlike PISA, NAEP does not administer different subject area assessments to the same students, and, therefore, no differentiation between main and minor domains is made. Rather, each student works on one domain. As with PISA, NAEP administers a questionnaire in addition to the subject-area testing, which aims to capture relevant contextual data for evaluating achievement results. These data are made available to the general public through a web-based application, the *NAEP Data Explorer* (NDE[4]). NAEP also captures teacher and school data via teacher and school administrator questionnaires and links these data to the student data. Compared to the international large-scale assessments, which include 30–35 min student questionnaires, NAEP has a considerably shorter student questionnaire of only 15 min. That is, the richness of the collected contextual data in the international assessments exceeds the richness that can be achieved with NAEP. Yet, NAEP benefits from keeping student burden low, thereby providing an optimal foundation for high participation rates and test-taker engagement.

[3] http://www.nationsreportcard.gov/

[4] http://nces.ed.gov/nationsreportcard/naepdata/

14.3 Psychosocial Skills in Large-Scale Assessments

By putting achievement results into context, questionnaire data from large-scale assessments can help educators and policymakers to better understand the circumstances under which learning and instruction take place; inform and shape school curriculum; support school-, district-, or state-level decisions in regard to the allocation of instructional resources; and help researchers further explore factors related to academic achievement. Figure 14.1 shows a schema proposed by Bertling (2014) to illustrate the different types of variables that are assessed in large-scale educational assessments. The schema specifically highlights the different categories of nonachievement data including psychosocial skills and noncognitive factors. On the right side of the figure, student achievement as the main outcome variable of policy interest is shown. The left-hand side of the figure depicts the different types of variables that are measured with the questionnaire component in large-scale assessments. Three main clusters can be differentiated within this set of variables.

The first cluster (labeled as "0" in Fig. 14.1) includes basic demographic and background variables, such as gender, age, or race/ethnicity. These variables provide relevant context for student achievement and often serve as grouping variables for score comparisons, but they do not play a key role in *explaining* the underlying reasons for differences in achievement scores.

The second cluster includes variables capturing factors related to differences in students' educational opportunities that might help describing and potentially explaining differences in students' achievement scores or achievement gaps.

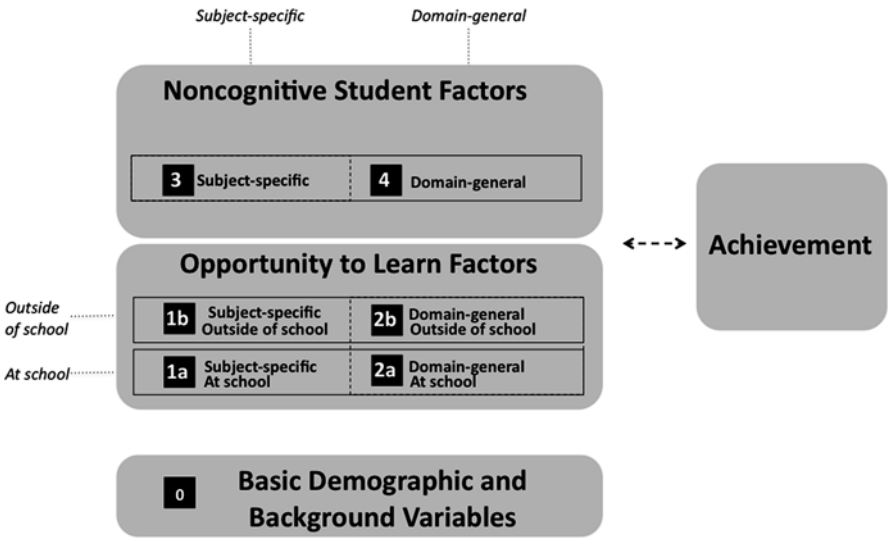

Fig. 14.1 Schematic model of types of variables collected with survey questionnaires in large-scale educational assessments; *left*, contextual variables; *right*, outcome variable

The term *opportunity to learn (OTL)* describes whether a student is exposed to possibilities to acquire knowledge and skills relevant to academic and workforce readiness and success. Early definitions described OTL quite narrowly as whether students had sufficient time to learn and received adequate instruction to learn in the specific subject of interest (Abedi, 2006; Carroll, 1963). Other works on the topic have expanded the concept. Researchers have suggested to define OTL not only based on subject-specific teacher instruction (Callahan, 2005; McDonnell, 1995), stressed the importance of quality of instruction in addition to mere quantity (Duncan & Murnane, 2011; Minor, Desimone, Spencer, & Phillips, 2015), pointed out the importance of informal learning and experiences in the home (Lareau & Weininger, 2003), and highlighted the need to evaluate OTL in country-specific contexts (Cogan & Schmidt, 2015). We define OTL as *all contextual factors that capture the cumulative learning opportunities a student was exposed to at the time of the assessment*. These factors comprise both learning opportunities at school and informal and formal learning outside of school. Examples of OTL at school include quantitative and qualitative aspects such as instruction time, exposure to relevant content, access to resources for learning, and presence of a positive school climate that promotes learning. Outside of school examples include time spent doing homework, reading for enjoyment, and use of computers to search for information. In addition, the *socioeconomic status* of a student's family (SES; White, 1982; for a definition see SES Expert Panel, 2012) and the academic climate at home can determine the student's learning opportunities (e.g., Lareau, 2011; Magnuson & Votruba-Drzal, 2009). For example, a student's exposure to relevant learning material, opportunities to practice, and received support might play an equal or even larger role for the student's success on a test than the student's general cognitive ability.

Figure 14.1 further distinguishes factors that are *subject specific* (such as learning time in a given subject or available instructional resources) and factors that are *domain-general* (such as technological equipment in the classroom or school climate). In total, these grouping variables allow distinguishing four different subclusters within the OTL cluster (labeled as 1a, 1b, 2a, and 2b).

OTL factors describe the learning opportunities students are exposed to that are, oftentimes, outside the direct control of the individual student. Variation in these factors is predominantly defined at the classroom, school, district, state, or country level or based on the students' socioeconomic status and family academic support. At the same time, students can decide whether or not to take chances as they come and actively seek out learning opportunities. Students differ in how they react to and approach learning situations. The strength of the associations between teaching and relevant learning outcomes also depends on whether and how students perceive the offered learning opportunities (Seidel & Reiss, 2014). This variation among students is captured with the variables in the third cluster, which includes psychosocial skills but is not limited to those.

Noncognitive student factors capture students' interest, motivation, self-related competency beliefs, attitudes toward school and learning, and other dispositions relevant to learning and achievement. Despite the importance of general cognitive ability and content knowledge to student achievement, over the past few decades,

educational, psychological, and econometric research has shown that psychosocial variables or "noncognitive factors" (Farrington et al., 2012; for an introduction of the term see Chap. 1; but also see Gehlbach, 2015a, or Kamenetz, 2015, for criticisms of the term "noncognitive") are of key importance for success in K-12 and beyond (Almlund, Duckworth, Heckman, & Kautz, 2011; Heckman & Kautz, 2013; Heckman, Stixrud, & Urzua, 2006; Nyhus & Pons, 2005; O'Connor & Paunonen, 2007; Paunonen & Ashton, 2001; Poropat, 2009; Richardson, Abraham, & Bond, 2012; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007), and relationships of certain noncognitive factors to achievement are comparable in strength to that of cognitive ability (Heckman & Kautz, 2013; Kautz et al., 2014; Poropat, 2009). The growing body of research on noncognitive factors suggests that success in school and beyond depends, for instance, on applying effort and being committed to succeed and persist during adversity, seeing learning as an opportunity, and respecting and understanding others. Additionally, noncognitive factors can interplay with cognitive factors in promoting a broad range of outcomes (Pekrun et al., 2007) and influencing career decisions (Parker et al., 2014). This view aligns with the definition of multidimensional goals of education, which emphasizes that teaching in schools should not only foster students' knowledge but also the development of personality, well-being, esthetics, and social competencies (e.g., Prenzel, 2012). For instance, for the domain of mathematics, the *National Council of Teachers of Mathematics* (1989) defined five goals in mathematics education with two of these goals (learn to value mathematics, become confident in one's ability to do math) representing noncognitive factors and three goals (become mathematics problem solvers, learn to communicate mathematically, learn to reason mathematically) representing content knowledge and skills (or cognitive factors).

In the context of large-scale educational assessments, measuring noncognitive factors matters because these constructs capture variables that may motivate a student to study harder, be more actively engaged in learning, and achieve higher grades, but, also in a broader sense, factors that make a student better prepared for adult life as a student and/or member of the workforce and an active citizen. While traditionally often seen as mostly stable person characteristics, more recent research findings (e.g., Farrington et al., 2012; Kautz et al., 2014) emphasize the malleability of noncognitive factors and the importance of teaching in fostering noncognitive components that help students become active learners who succeed in school. Several studies have pointed to the specific importance of teachers' behaviors such as setting grades, providing constructive feedback, and promoting mastery experiences, especially at early grades (Lent & Brown, 2006; Stock & Cervone, 1990). In a recent study based on PISA 2012 data, Schiepe-Tiska, Heine, Luedtke, Seidel, and Prenzel (2015) showed that cognitive activation (Lipowsky et al., 2009; see also Hiebert & Grouws, 2007), classroom management (Kunter, Baumert, & Köller, 2007), and student support (Pintrich, Marx, & Boyle, 1993) as the three basic dimensions of instructional quality show distinct patterns of relationships with multidimensional learning outcomes, both at the individual and at the classroom level. Findings from individual differences researchers further provide ample validity evidence for the malleability, amenability for interventions, and lifetime changes of

noncognitive factors (see Chap. 4; Abraham, 2012; Ashford, Edmunds, & French, 2010; Bandura, 1994; Dweck, Walton, & Cohen, 2011; Heckman & Kautz, 2013; Richardson et al., 2012; Specht, Egloff, & Schmukle, 2011).

14.4 The Shifting Role of Variables Measured with Questionnaires in Large-Scale Assessments

Since the first introduction of large-scale educational assessments, the focus of educators, policymakers, and the general public has broadened from relatively few content domains to a larger set of domains including new skills such as collaborative problem solving, technology and engineering literacy, and financial literacy. In parallel with this broadened interest, the interest in variables gathered through the contextual questionnaires of these assessments has also increased and broadened. Questionnaires started to focus more heavily on domain-general factors rather than focusing on narrow variables relevant to specific subject areas only. Figure 14.2 illustrates this trend.

This change also reflects the shift from viewing these variables primarily as contextual factors that serve to better understand cognitive scores to viewing noncognitive variables as measures of separate constructs of their own interest that might be interpreted as important achievement predictors and represent additional outcomes of education (Ginsburg & Smith, 2013; Rychen & Salganik, 2003). This is reflected in trends to feature noncognitive variables more prominently in reports – for the most recent PISA cycle, only one report focuses on achievement results alone, while several reports of comparable length dive into the relationships with contextual variables or country differences in noncognitive or OTL factors. Efforts have increasingly focused on presenting results as outcomes ("Noncognitive Outcomes"; OECD, 2013b), investigating trends (OECD, 2014b), comparing subgroups (OECD, 2014a), and considering the constructs measured with questionnaires for use in teacher accountability metrics (e.g., Gabrieli, 2015). Additionally, recent theoretical and methodological developments have pushed for the adoption of measures of noncognitive skills when assessing the efficiency of different policy interventions (Stiglitz, Sen, & Fitoussi, 2010).

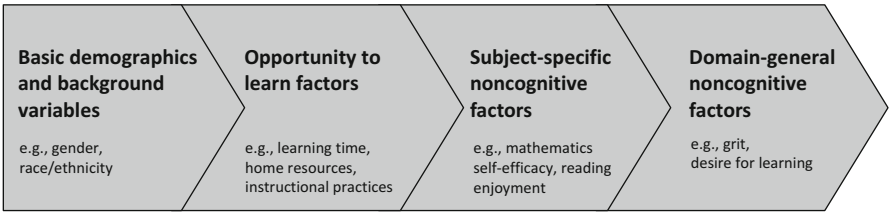

Fig. 14.2 Transition from basic demographics to domain-general noncognitive factors

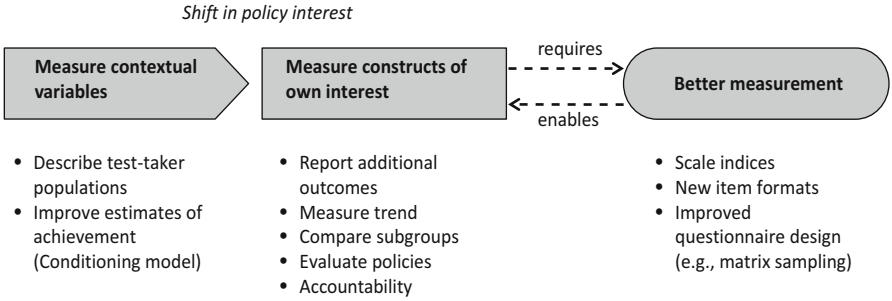

Fig. 14.3 Shift in policy interest and the role of better measurement

Figure 14.3 illustrates the shift in policy interest from contextual variables to constructs of their own interest and the challenges associated with this shift.

Governing bodies of major large-scale assessments have called for more noncognitive measurement and promoted subgroup comparisons and rankings based on psychosocial or noncognitive factors. In the United States, Smith et al. (2012) described the NAEP contextual variables as "an underused national resource" (p. 1) calling for making greater use of noncognitive data for reporting and policy. The *National Assessment Governing Board*'s first policy principle in their *Policy Statement on NAEP Background Questions and the Use of Contextual Data in NAEP Reporting* explicitly highlights the importance of "non-cognitive questions asked of students, teachers, and schools" for enriched NAEP reporting (National Assessment Governing Board, 2012, p. 1). An example for concerted efforts in emerging countries aimed at large-scale noncognitive measurement is the *social and emotional development and school learning* project in Brazil (Santos & Primi, 2014) planned by the Ayrton Senna Institute in partnership with the Rio de Janeiro State Education Department (SEEDUC) and OECD.

Since the early 2000s, when PISA was first introduced, questionnaires started to increasingly capture domain-general factors that are not directly linked to a particular subject. The PISA 2015 Assessment Framework now explicitly includes a high-priority module "domain-general student attitudes and behaviors," which aims to measure personality facets, well-being, time use, and health (Bertling & Kyllonen, 2012; OECD, 2013f). The same trend can be observed in NAEP where the scope of variables measured with the student, teacher, and school questionnaires has been broadened from mostly subject-specific factors to core contextual factors cutting across all subject areas. For the 2017 NAEP assessment cycle, NCES started developing noncognitive modules including technology use, school climate, grit, and desire for learning (Bertling, 2014).

Constant evolution is necessary for large-scale surveys given that education in the United States and worldwide is undergoing continual changes which, by default, means that the information that is most relevant and useful for teachers, districts, states, policymakers, and the general public is constantly changing as well.

14.5 Education Policies that Can Help Promote Psychosocial Skills

The inclusion of noncognitive factors in large-scale international surveys of K-12 populations can lead to the identification of policies and practices that promote student overall development, not only academic achievement. Under the assumption that student samples are representative of the students of each country, country-level differences in both subject-specific and psychosocial skills can be considered to be the result of the opportunities and experiences students gain, because of formal and informal learning, that occurs with the family, at school, and in the neighborhood, as well as education policies, social norms, and practices (cf. Fig. 14.1). Therefore, large-scale international assessments can shed light on the factors that are most strongly associated, either positively or negatively, with student development. Most of the analyses that are conducted using large-scale international assessment data are descriptive and are not designed to establish causal relations. However, in a number of circumstances, differences in the timing and the features of particular education policies – in different countries or regions within countries – can be used to reliably evaluate the effect of specific policies on student outcomes (Agasisti & Cordero-Ferrera, 2013; Ciccone & Garcia-Fontes, 2009; Jakubowski, Patrinos, Porta, & Wisniewski, 2010). Moreover, the fact that large-scale international assessments such as TIMMS, PIRLS, PISA, and NAEP have been repeated at regular intervals over the past few decades means that they can be used as pseudo-panels to evaluate trends and factors that help to explain temporal variations in achievement and psychosocial and emotional skills (OECD, 2013a, 2013b).

Information arising from student self-reports in the questionnaire portion of large-scale international assessments can be complemented by information gathered through questionnaires distributed to the students' teachers or their school principals or by aggregating student responses from the same school or class (OECD, 2013a, 2013b, 2013e). Moreover, student-level data can be complemented by system-level information on regional and nationwide policies. Such data can be collected either through desk-based research by individual scientists or through coordinated system-level data collection exercises aimed at gathering information in ways that guarantee the accuracy and cross-national validity of estimates (OECD, 2014a; UNESCO, 2014).

Examining the interplay between cognitive and noncognitive factors is essential to better understand what policies can promote learning and equip students to face the challenges that lie ahead. In the following, we illustrate, based on the example of horizontal stratification, how educational policies may influence student noncognitive factors. We then give an example from one country, Japan, how educational policies might address the importance of noncognitive factors for student learning.

The essential features of how educational systems are organized also reflect assumptions about how students learn, when they thrive, and what motivates them. For example, school systems face the challenge of providing instruction to student populations that are increasingly diverse in terms of socioeconomic and demographic

profiles, abilities, aspirations, and expectations. To meet these challenges, some school systems (e.g., Australia, Finland, Poland, or the United States; OECD, 2011) use nonselective and comprehensive policies that seek to provide all students with similar opportunities, leaving it to each teacher and school to cater to the full range of student abilities, interests, and backgrounds. Other systems respond to diversity by grouping students between schools or classes within schools to instruct students according to their academic potential and/or interests in specific programs (e.g., Germany, Hungary, Singapore, or the Netherlands; OECD, 2011; but note recent developments to reduce stratification and better integrate different learners in Germany; Maaz et al., 2013). Horizontal stratification refers to the practice of differentiating instruction within a specific grade or education level creating groups of students according to their interests and/or performance. School systems, for example, can offer specific programs (such as vocational or academic programs), define the age at which students are admitted into these programs, and determine the extent to which academic records are used to select students for their schools. Teaching in these schools or classes may be adapted to students with different needs, with class sizes and teacher assignments determined accordingly. How schools are organized and what stratification strategy is applied might influence student motivation and learning. PISA 2012 findings suggest associations between the way in which students are assigned to programs or schools and their level of instrumental motivation to learn mathematics. In education systems that adopt horizontal stratification practices, students' performance is more strongly associated with the socioeconomic condition of their family than it is the case elsewhere (OECD, 2013c). Moreover, there is a strong negative association between the levels of students' motivation and the degree to which systems sort and group students into different schools and/or programs (OECD, 2013c). In those systems that tend to separate students into different schools or programs, students generally report lower levels of instrumental motivation to learn mathematics than students who are in systems with less between-school horizontal stratification (OECD, 2013c). These findings from PISA mirror previous reports on the role of organizational features in promoting motivation, aspirations, and goal setting among secondary school children (Buchmann & Dalton, 2002; Buchmann & Park, 2009; Monseur & Lafontaine, 2012). Motivation also tends to be lower in countries in which larger proportions of students attend vocational or prevocational rather than academic programs or where a large proportion of students attends academically selective schools (e.g., the Netherlands or the Czech Republic, respectively, as compared to Canada, the United Kingdom, or the United States). Motivation appears to be lower also in the presence of policies that create distinct student groups or select students early for different academic trajectories, rather than at a later age, and deal with behavioral or academic problems by enabling schools to transfer such students to other schools (OECD, 2013c).

Selecting and sorting students generally acts as an indirect form of segregation that may reinforce socioeconomic disparities, result in differences in opportunities to learn, and potentially demotivate considerable proportions of students who do not feel that the system believes in their potential. At the same time, education systems may view grouping students as a way to promote learning by catering to the

specific aspirations and ability of different students. However, because students' abilities and what motivates and interests them is very malleable during childhood and teenage years (Gorey, 2001; Heine et al., 2001) with noncognitive factors being more malleable than cognitive ability (e.g., Kautz et al., 2014), horizontal stratification policies run the risk of demotivating the very students that would need a system that sets high aspirations for them. Recent research suggests that tracking type may moderate adverse consequences on students' motivation and self-beliefs (Chmielewski, Dumont, & Trautwein, 2013). This further supports the importance of a strong theoretical and empirical knowledge base upon which research on noncognitive skills in large-scale assessments could be anchored. Such research could, in turn, inform the development of directions in education policy and balance the objectives of promoting students' development of strong cognitive and noncognitive skills.

14.5.1 Japan: Reforming Curricula to Consider the Importance of Psychosocial Skills

Japan is a country that has used new evidence emerging from large-scale assessments of learning outcomes to reform its curriculum and emphasize the importance of a comprehensive approach to education. Japanese students are consistently among the top performers in PISA. Japan's mean score in reading in both 2000 (520 points) and 2009 (522 points) placed the country among the top ten performers in the world. Between PISA 2009 and PISA 2012, average reading scores increased to 538 points, showing that improvements are possible even among top-performing countries. In science, similar improvement was observed between PISA 2006 and PISA 2012 as average scores improved from 531 to 547 points, at an average rate of 2.6 points per year. Despite these high levels of performance, PISA results prompted wide discussion of policy reforms to offer equal opportunities to all children and a curriculum appropriate for the twenty-first century (OECD, 2012). In 2006, Japan amended the Basic Act on Education, which had regulated education services for the previous 60 years. The amendment modified the legal framework, stated objectives of education, introduced a system for renewing educational personnel certificates, and revised the administration of local education authorities to improve the role of the local boards of education. These changes implied moving toward an education model that emphasized a good balance between cognitive and noncognitive knowledge and skills (OECD, 2012).

Previous PISA results highlighted comparatively low levels of engagement, motivation, and mathematics self-beliefs among Japanese students. When compared to students in 2003, however, Japanese students in 2012 reported a stronger sense of belonging, lower rates of tardiness, better attitudes toward school, and higher levels of intrinsic and instrumental motivation to learn mathematics. In the past decade, the student experience and the relationship between schools and the community also changed. School-community cooperation has become indispensable. For example, parents and community members now take some responsibilities for

managing schools and help in teachers' lessons to encourage the connection between school and the outside world. Coincidentally, students in 2009 were more likely to enjoy reading and perform better in open-ended constructed tasks in PISA than their counterparts in 2000 (OECD, 2010). The Japanese also became concerned that their continued dominance in generating and exploiting advanced technologies was under threat. While Japanese students continued to do well in international comparisons of achievement in mathematics and science, results seemed to indicate they enjoyed science less than students elsewhere as they progressed through schooling. These findings were also reflected in results from the PISA 2003 mathematics assessment, which showed high levels of student anxiety about mathematics and low levels of interest in and enjoyment of the subject. Concerns about a loss in moral standards and declining student motivation coincided with a perceived decline in Japan's edge in innovation. In 1996, the Japanese Ministry of Education, Culture, Sports, Science, and Technology began to apply a new philosophy to education that was intended to enhance students' ability to act autonomously and think creatively. *Ikiru chikara*, or "zest for living," emphasized key competencies, independent thinking, and problem-solving skills. A key part of this reform was to set the conditions that would enable students to develop a well-rounded personality and promote the development of the cognitive and noncognitive competencies that are needed in the country's changed economy and society.

14.6 Challenges and the Promise of Better Measurement

While the policy interest in measuring noncognitive factors in large-scale assessments has increased, the time allocated for students to fill out respective questionnaires has stayed virtually constant. The quality of the data gathered through questionnaires relies on students being willing to engage with the material and provide accurate information. Research suggests that in the presence of long questionnaires, data quality suffers (Galesic & Bosnjak, 2009). Therefore, the goal of broadening the spectrum of outcomes being measured is challenged by the need to keep the burden placed on participating students low to ensure high-quality data.

In addition to time constraints, there are a number of conceptual and methodological challenges in mapping students' psychosocial and emotional skills in national and international contexts. Even more difficult is identifying the role that education systems, individual schools, and families play in shaping students' social and emotional growth. There is, for example, no clear definition of, or consensus on, what education systems should strive to develop beyond subject-specific skills (Forster, 2004). Although the aim of large-scale educational assessments has broadened to capture behavioral differences and differences in students' psychosocial and emotional skills, these assessments largely rely on self-reported answers gathered from students, school principals, teachers, and, in some countries, students' parents. What these studies capture might, therefore, be considered a mixture of differences in students' behaviors, beliefs, attitudes, and perceived desirability of certain

response options of the administered questions. Moreover, a set of behaviors, beliefs, and attitudes can be more or less appropriate depending on cultural contexts. For example, in some countries, teenagers that express a high degree of interest in school subjects and spend considerable time studying are socially sanctioned by their peers; in others, such students command respect and are viewed as positive examples by other students (Ladd, Kochenderfer-Ladd, Visconti, & Ettekal, 2012).

Large-scale international assessments are limited in scope by the need to examine constructs in a comparative perspective, meaning that differences in results across countries on self-reports might reflect how students in different countries understand the wording of specific questions, may be differently responsive to social desirability and acquiescence, or exhibit differences in the use of response scale categories. These differences may be responsible for what is referred to as the "attitude-achievement paradox" (Mickelson, 1990; Van de Gaer, Grisay, Schulz, & Gebhardt, 2012) whereby relationships between noncognitive factors and achievement are positive within countries but are negative across countries. In the context of PISA, inconsistencies in the directionality of relationships between constructs measured with Likert scales (e.g., attitude toward mathematics) and achievement outcomes (e.g., mathematics proficiency), on the individual versus aggregated level, were found consistently across past assessment cycles (Bertling & Kyllonen, 2013; Kyllonen & Bertling, 2013; OECD, 2010; Van de Gaer et al., 2012). The relationships at the within-country level (e.g., positive attitudes relate to higher proficiency) go in line with underlying psychological theories about the relationship between attitudes, motivation, and behavior (Atkinson, 1957; Eccles et al., 1983, 1993). At the same time, no construct-specific theoretical models for the explanation of country-level relationships are available (Mõttus et al., 2012). As described in Kyllonen and Bertling (2013), such inconsistencies might stem from real differences in how relationships play out at the individual and country level or be the result of systematic differences between countries in how students interpret the agreement response scale or response styles (Buckley, 2009; Cheung & Rensvold, 2000). Comparative conclusions might further be threatened by the lack of measurement invariance across groups (Milfont & Fischer, 2010). If the issue of cross-cultural and subgroup differences in survey response styles is not considered, and existing response styles are not corrected for, secondary analysts who use noncognitive data are at risk of reaching erroneous conclusions, especially as they draw inferences to the individual level based on aggregated data ("ecological fallacy"; Freedman, 1999; Robinson, 1950). Stakeholders might, for instance, wrongly infer that more positive attitudes toward mathematics lead to lower levels of mathematics achievement based on inspecting country average scores in mathematics achievement and Likert-based indices of mathematics motivation in PISA 2012 (OECD, 2013b, 2013d). Inappropriate or insufficient measurement of noncognitive factors can seriously threaten the validity of claims derived from the resulting data, potentially limiting the usefulness of these data for high-stakes policy decisions (Duckworth & Yeager, 2015).

International and national large-scale assessments have started exploring new directions to improve the measurement of noncognitive and OTL factors to enable

policymakers and educators to make evidence-based decisions. Three promising directions in improved measurement are presented in the following section. These include the use of new item formats, improved questionnaire pretesting, and new questionnaire designs in technology-based environments.

14.6.1 New Item Formats

Despite the large body of research demonstrating the problematic nature of Likert-type agree/disagree items (e.g., Converse & Presser, 1986; Dillman, Smyth, & Christian, 2014; Fowler, 2009), the majority of large-scale assessments currently use traditional multiple choice and matrix item formats (i.e., close-ended questions that ask participants to provide responses in one or more rows using the same set of response options). In order to compensate for the limitations described above, large-scale assessments have recently begun introducing new and innovative item formats to replace the 4-point Likert scale format with vague response categories such as "strongly agree," "agree," "disagree," and "strongly disagree." PISA 2012 introduced several new item formats for increased cross-cultural validity of the derived questionnaire indices, among them anchoring vignettes to adjust Likert-type responses (Bertling & Kyllonen, 2013), topic familiarity items with overclaiming correction (Kyllonen & Bertling, 2013), and situational judgment tests (SJT; Weekley & Ployhart, 2006) to measure students' problem-solving approaches (Bertling, 2012). Anchoring vignettes (Hopkins & King, 2010; King & Wand, 2007) are short descriptions of hypothetical individuals in a particular situation that provide a basis for individual-level adjustment of responses to account for a respondent's interpretation of the response scale. The anchoring vignettes approach has been used for making cross-country comparisons in various fields of research (Kapteyn, Smith, & Van Soest, 2007; Kristensen & Johansson, 2008; Salomon, Tandon, & Murray, 2004). Although the method has been used in sociological, political science, and other research, it had not been used in large-scale educational assessments prior to PISA 2012. Items with overclaiming correction (Ackerman & Ellingsen, 2014; Feeney & Goffin, 2015; Paulhus & Dubois, 2014; Paulhus, Harms, Bruce, & Lysy, 2003) allow for improved measurement of topic familiarity by accounting for respondents' tendencies to overstate their perceived knowledge in certain areas. SJTs contextualize survey questions within scenarios, thereby potentially reducing general response tendencies, effects of socially desirable responding, faking, or adverse impact (McDaniel, Hartman, Whetzel, & Grubb, 2007; Oostrom, De Soete, & Lievens, 2015; Whetzel & McDaniel, 2009).

Kyllonen and Bertling (2013) provided an overview of new item types in PISA 2012 that led to higher validity of the noncognitive instruments used in the study. Figures 14.4, 14.5, and 14.6 show examples of the innovative item formats used in PISA 2012, specifically anchoring vignettes (Fig. 14.4), topic familiarity items with overclaiming correction (Fig. 14.5), and SJTs (Fig. 14.6). The anchoring vignettes shown follow one of several possible development approaches. Kyllonen and

ST82

Q **Below you will find descriptions of three mathematics teachers. Read each of the descriptions of these teachers, then let us know to what extent you agree with the final statement.**
(Please tick only one box in each row.)

	Strongly agree	*Agree*	*Disagree*	*Strongly disagree*
a) Ms. <name> sets mathematics homework every other day. She always gets the answers back to students before examinations.**Ms. <name> is concerned about her students' learning.**	■1	■2	■3	■4
b) Mr. <name> sets mathematics homework once a week. He always gets the answers back to students before examinations.**Mr. <name> is concerned about his students' learning.**	■1	■2	■3	■4
c) Ms. <name> sets mathematics homework once a week. She never gets the answers back to students before examinations.**Ms. <name> is concerned about her students' learning.**	■1	■2	■3	■4

Fig. 14.4 Example for innovative item formats used in PISA 2012: anchoring vignettes (OECD, 2013b)

Bertling (2015) presented a classification of different possible approaches for developing anchoring vignettes. The overclaiming items shown here were developed to improve the measurement of familiarity with mathematics concepts in the context of PISA 2012. The technique is not bound to the area of mathematics but can be extended into other subject areas as well (e.g., Ziegler, Kemper, & Rammstedt, 2015). The SJT item shown here is one of a set of three scenarios used in PISA 2012 to measure students' approaches to problem solving (see OECD, 2013b, for the other two scenarios included). Different types of problem-solving behaviors could be distinguished based on students' ratings of different possible behaviors under each scenario (Bertling, 2012).

NAEP introduced several new item formats in their 2015 *Computer Access and Familiarity Study* (CAFS) and is further exploring these item types for future NAEP cycles. The CAFS focused on identifying the technology available to students both in and outside of school, how technology is used in the classroom, along with students' familiarity with specific computer tasks (e.g., typing, looking up information on the Internet, editing text, and maintaining a website or a blog). In addition to traditional self-report items, the CAFS questionnaire included anchoring vignettes

ST62

Q **Thinking about mathematical concepts: how familiar are you with the following terms?**

(Please tick only one box in each row.)

	Never heard of it	*Heard of it once or twice*	*Heard of it a few times*	*Heard of it often*	*Know it well, understand the concept*
a) Exponential Function	■1	■2	■3	■4	■5
b) Divisor	■1	■2	■3	■4	■5
c) Quadratic Function	■1	■2	■3	■4	■5
d) Proper Number*	■1	■2	■3	■4	■5
e) Linear Equation	■1	■2	■3	■4	■5
f) Vectors	■1	■2	■3	■4	■5
g) Complex Number	■1	■2	■3	■4	■5
h) Rational Number	■1	■2	■3	■4	■5
i) Radicals	■1	■2	■3	■4	■5
j) Subjunctive Scaling*	■1	■2	■3	■4	■5
k) Polygon	■1	■2	■3	■4	■5
l) Declarative Fraction*	■1	■2	■3	■4	■5
m) Congruent Figure	■1	■2	■3	■4	■5
n) Cosine	■1	■2	■3	■4	■5
o) Arithmetic Mean	■1	■2	■3	■4	■5
p) Probability	■1	■2	■3	■4	■5

Fig. 14.5 Example for innovative item formats used in PISA 2012: topic familiarity with overclaiming correction (OECD, 2013b)

pertaining to computer familiarity and implemented the overclaiming technique for enhanced measurement of students' familiarity with technology-related terms (Bertling & Almonte, 2014). Overclaiming items combined computer-related terms with terms from neuroscience to create foils that might serve as reference points in rescaling students' subjective familiarity ratings.

Additionally, the transition of student questionnaires from paper administration to administration on laptops or tablets allows for the use of soft edits and consistency checks that can make previously unfeasible response formats, such as free

ST101

Q *Suppose that you are planning a trip to the zoo with your brother. You don't know which route to take to get there.*

What would you do? For each suggestion, tick the option that best applies to you.

(Please tick only one box in each row.)

	I would definitely do this	*I would probably do this*	*I would probably not do this*	*I would definitely not do this*
a) I read the zoo brochure to see if it says how to get there.	■1	■2	■3	■4
b) I study a map and work out the best route.	■1	■2	■3	■4
c) I leave it to my brother to worry about how to get there.	■1	■2	■3	■4
d) I know roughly where it is, so I suggest we just start driving.	■1	■2	■3	■4

Fig. 14.6 Example for innovative item formats used in PISA 2012: situational judgment test (OECD, 2013b)

text entry, viable replacements for close-format multiple choice items. Given the shift in education and large-scale assessments from paper-and-pencil assessments to digital assessments, affordances can be included to collect information that is more useful and provide a better user experience. For instance, skip patterns are predetermined designs in which an item or items are bypassed based on a response provided by the participant to a particular previously answered item. Collection of response time allows for creation of time effort indicators which allow distinguishing between solution behavior and rapid-guessing behavior as each item response can be classified as one of these two behaviors by comparing the response time to a predetermined time threshold established for that item (e.g., Wise & Kong, 2005). Digital assessments might also lead a way to the development of behavioral measures of certain noncognitive skills, such as perseverance and effort, through the analysis of process data. The use of process data in the context of large-scale educational assessments represents a fertile area of future work that bears the promise of delivering new empirical insights on students' noncognitive factors. Keehner et al. (2014) presented an example for the types of indicators beyond classical achievement scores that can be derived based on process data from large-scale assessments. In the context of the 2014 NAEP Technology and Engineering Literacy (TEL) assessment, indicators of students' levels of efficiency and systematicity during problem solving could be created based on students' behaviors while working on a scenario-based interactive computer task. Using process data and behavioral measures creates an opportunity to circumvent the limitations of self-reported responses. Yet, students' behavior is sampled on very narrowly defined tasks in very

specific situations. Therefore, advances in the understanding of students' psychosocial skills will likely rely on a combination of two approaches – the examination of self-reports and of behaviors.

14.6.2 Improved Questionnaire Pretesting

In order to make best use of the limited assessment time available to collect contextual information, questionnaire pretesting gains even more importance than with smaller lab studies. Systematic qualitative pretesting of questionnaire items with small samples through cognitive interviewing can help determine optimal item formats and identify problems and respective solutions before moving into costly large-scale pilot studies. Experimentation with item wording and item formats, for instance, can take place during pretesting activities such as cognitive interviews and focus groups. In NAEP, all newly developed student, teacher, and school administrator survey items go through rigorous reviews and are pretested in cognitive interviews before any pilot and operational administrations. Cognitive interviews are conducted as structured one-on-one interviews with trained interviewers gaging students' understanding and answer processes. In preparation for the 2017 operational administration, cognitive interviews addressed a variety of research questions, such as comparing item wording for self-efficacy items (e.g., "Do you think you would be able to do each of the following?" versus "How confident do you feel that you could do each of the following?") and contrasting four-point, five-point, or six-point scales to determine the best wording for students from all geographic, cultural, and socioeconomic backgrounds (Alegre, Almonte, Anthony, & Bertling, 2015). As part of the NAEP cognitive interview pretesting, activity item formats were compared, specifically multiple choice versus free response, to determine the best way of capturing information such as classroom size, number of computers available in the school, frequency of being assigned writing assignments, and percentage of time spent on specific classroom activities (Alegre et al., 2015). NAEP recently introduced eye tracking as another innovative method for questionnaire pretesting allowing for in-depth and noninvasive capture of students' actions in answering survey items on tablets. Findings from a recent eye-tracking study allowed insights into how fourth-grade students interacted with various item types on tablets (Maitland, Sun, Tourangeau, Almonte, & Bertling, 2015). A key finding from this study was that fourth-grade students answered both simple discrete and more complex survey items following the intended sequence of processing steps and that students enjoyed answering matrix-type questions more than stand-alone discrete questions. Implementing thoroughly designed pretesting procedures for questionnaires to be used in large-scale assessments helps ensure that data from large-scale pilots and operational administrations can be meaningfully interpreted. This research phase therefore constitutes a crucial stage in the development of theoretically valid and practically useful instruments, especially when questionnaire items for diverse populations are developed.

14.6.3 Matrix Sampling for More Efficient Questionnaire Designs

Facing the constraints in overall testing time and the large sample sizes in PISA, TIMSS, and NAEP, enhanced questionnaire designs with questions being spiraled across multiple forms become a viable option to reduce burden while maintaining content coverage across relevant areas. Matrix sampling approaches are the standard practice for the subject-area tests in large-scale educational assessments (Comber & Keeves, 1973; OECD, 2013b). PISA as the first large-scale assessment made use of a matrix sampling design also for the student questionnaire in the 2012 assessment cycle that allowed for an increase of questionnaire content by 33 % (OECD, 2014b). In the context of NAEP, the National Assessment Governing Board recommended that "whenever feasible, assessment samples should be divided (spiral sampling … in order to cover more topics without increasing respondent burden" (National Assessment Governing Board, 2012, p. 3). While research findings to date mostly point to the benefits of questionnaire spiraling in the context of large-scale assessments, methodological concerns have been raised as well. On the one hand, findings indicate that spiraling allows for substantial increase in content coverage with very small to negligible impact on the overall measurement model, including conditioning and estimation of plausible values (Adams, Lietz, & Berezner, 2013; Almonte, McCullough, Lei, & Bertling, 2014; Kaplan & Wu, 2014; Monseur & Bertling, 2014). On the other hand, spiraling might cause biases in estimation results depending on what spiraling designs are used and what matrix sampling designs are implemented for the cognitive assessment (von Davier, 2014). Further research on different matrix sampling approaches for questionnaires is needed to determine the most feasible approach for future implementation.

Regardless of what specific spiraling approach is considered, a matrix sampling design cannot increase the amount of information that is collected at the individual student level and might, depending on the specific matrix sampling design, limit the possibility of conducting multivariate analyses. Explorations of potential matrix sampling in future large-scale assessments should therefore focus on possible use of matrix sampling in conjunction with imputation approaches (Kaplan & Wu, 2014). The quality of results derived from models based on imputed datasets may also depend on the specific assumptions adopted when imputing and on the spiraling design of the study. For instance, spiraling designs might divide content into blocks with the aim of keeping scales intact and maximizing overlap between closely related constructs (such as in PISA 2012) or separate items within scales into different blocks, thereby ensuring that every student answers a subset of items for each construct.

14.7 Summary and Conclusions

National and international large-scale assessments no longer focus on achievement results as their only key outcome. Programs have shifted attention to also report on noncognitive student factors relevant to achievement and general life outcomes as

independent or additional reporting elements. This chapter described which types of psychosocial or noncognitive factors are measured with student questionnaires in these assessments for K-12 populations, internationally and within the United States. PISA and NAEP currently take the roles of the largest international and national assessment programs, respectively, and both devote increased attention to the measurement of psychosocial skills and noncognitive student factors in addition to measuring what students know and can do in several subject domains. We described a schematic model (see Fig. 14.1) that can serve as a helpful guide in distinguishing different kinds of variables measured in large-scale assessments. Basic demographic and background variables can be distinguished from opportunity-to-learn factors and noncognitive student factors, the latter representing the cluster of constructs that this volume mainly focuses on.

In parallel with the broadened interest of policymakers and educators in constructs that lie beyond literacy in a traditional sense, the interest in variables gathered through questionnaires filled out by students, teachers, or school administrators has increased substantially and started to focus more heavily on domain-general factors rather than narrow variables relevant to specific subject areas. This change reflects the shift from perceiving these variables primarily as achievement predictors to viewing them also as stand-alone measures of constructs of their own interest. Examples of this shift are the inclusion of personality facets in PISA 2012 (OECD, 2013b) and the introduction of constructs such as grit and desire for learning to NAEP (Bertling, 2014). The increased attention to noncognitive student factors, and the commitment to report on additional outcomes, measure trends, compare subgroups, and provide a relevant database to evaluate policies, poses new demands on questionnaires in large-scale assessments. Current measurement approaches rely heavily on Likert-type self-report items that are prone to social desirability or cross-cultural differences and reference group effects that might compromise the validity of conclusions drawn from the data (Duckworth & Yeager, 2015; Gehlbach, 2015b; Kyllonen & Bertling, 2013).

Three promising directions in improved measurement targeted at increased validity and subgroup comparability under the constraint of keeping student burden low are the use of new item formats, improved questionnaire pretesting processes, and new questionnaire designs for technology-based environments. We will likely see additional new trends over the coming years as the measurement of psychosocial skills in large-scale assessments continues to evolve. In a world where change is the only constant and one of the core goals for education systems becomes promoting lifelong learners who are able as well as eager to face the demands and challenges of a truly global society, large-scale assessments need to broaden their focus to stay relevant. Their success as helpful tools and monitoring systems for policymakers and educators will depend to a growing extent on how well they can serve the policy and public demand for enhanced noncognitive measurement.

References

Abedi, J. (2006). Psychometric issues in the ELL assessment and special education eligibility. *The Teachers College Record, 108*, 2282–2303.

Abraham, C. (2012). Mapping change mechanisms and behavior change techniques: A systematic approach to promoting behavior change through text. In C. Abraham & M. Kools (Eds.), *Writing health communication: An evidence-based guide for professionals* (pp. 99–116). London: Sage.

Ackerman, P. L., & Ellingsen, V. J. (2014). Vocabulary overclaiming. A complete approach: Ability, personality, self-concept correlates, and gender differences. *Intelligence, 46*, 2016–2227.

Adams, R. J., Lietz, P., & Berezner, A. (2013). On the use of rotated context questionnaires in conjunction with multilevel item response models. *Large-Scale Assessments in Education, 1*(1), 1–27.

Agasisti, T., & Cordero-Ferrera, J. M. (2013). Educational disparities across regions: A multilevel analysis for Italy and Spain. *Journal of Policy Modeling, 35*(6), 1079–1102.

Alegre, J. M., Almonte, D. E., Anthony, J. S., & Bertling, J. P. (2015, March). *Key findings from NAEP survey questionnaire cognitive interviews.* Presentation at the NAEP Questionnaire Standing Committee Meeting, Washington, DC.

Almlund, M., Duckworth, A., Heckman, J. J., & Kautz, T. (2011). *Personality psychology and economics* (no. w16822). National Bureau of Economic Research.

Almonte, D. E., McCullough, J., Lei, M., & Bertling, J. P. (2014, April). Spiraling of contextual questionnaires in the NAEP TEL pilot assessment. In J. P. Bertling (Chair), *Spiraling contextual questionnaires in educational large-scale assessments*. Symposium conducted at the meeting of the National Council on Measurement in Education, Philadelphia.

Ashford, S., Edmunds, J., & French, D. P. (2010). What is the best way to change self-efficacy to promote lifestyle and recreational physical activity? A systematic review with meta-analysis. *British Journal of Health Psychology, 15*(2), 265–288.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64*(6, Pt. 1), 359–372.

Bandura, A. (1994). *Self-efficacy*. New York: Wiley.

Bertling, J. P. (2012, October). *Students' approaches to problem solving-comparison of MTMM models for SJT data from PISA 2012*. Technical report presented at the PISA 2012 Questionnaire Expert Group and Problem Solving Expert Group Meeting, Heidelberg, Germany.

Bertling, J. P. (2014, May). *Improving the contextual questionnaires for the National Assessment of Educational Progress: Plans for NAEP Core contextual modules*. White paper presented at the meeting of the National Assessment Governing Board, Boston.

Bertling, J. P., & Almonte, D. E. (2014). *Anchoring vignettes to improve the measurement of students' familiarity with computers and digital technology in the National Assessment of Educational Progress (Internal technical memo)*. Princeton, NJ: Educational Testing Service.

Bertling, J. P., & Kyllonen, P. C. (2012). *Domain-general student attitudes and behaviors: Constructs and items for PISA 2015 (Module 10).* Technical report presented at PISA 2012 Questionnaire Expert Group meeting, Heidelberg, Germany.

Bertling, J. P., & Kyllonen, P. C. (2013, August). Using anchoring vignettes to detect and correct for response styles in PISA questionnaires. In M. Prenzel (Chair), *The attitudes-achievement-paradox: How to interpret correlational patterns in cross-cultural studies*. Symposium conducted at the EARLI 2013 conference, Munich, Germany.

Buchmann, C., & Dalton, B. (2002). Interpersonal influences and educational aspirations in 12 countries: The importance of institutional context. *Sociology of Education, 75*, 99–122.

Buchmann, C., & Park, H. (2009). Stratification and the formation of expectations in highly differentiated educational system. *Research in Social Stratification and Mobility, 27*(4), 245–267.

Buckley, J. (2009). *Cross-national response styles in international educational assessments: Evidence from PISA 2006*. Retrieved from https://edsurveys.rti.org/PISA/documents/Buckley_PISAresponsestyle.pdf

Callahan, R. M. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal, 42*(2), 305–328.

Carroll, J. (1963). A model of school learning. *The Teachers College Record, 64*(8), 723–733.

Cheung, G. W., & Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross-Cultural Psychology, 31*, 187–212.

Cogan, L. S., & Schmidt, W. H. (2015). The Concept of Opportunity to Learn (OTL). In K. Stacey & R. Turner (Eds.), *International Comparisons of Education. In Assessing Mathematical Literacy* (pp. 207–216). Cham: Springer International Publishing.

Chmielewski, K., Dumont, H., & Trautwein, U. (2013). Tracking effects depend on tracking type: An international comparison of academic self-concept. *American Educational Research Journal, 50*, 925–957.

Christenson, S., Reschly, A. L., & Wylie, C. (2012). *Handbook of research on student engagement*. New York: Springer.

Ciccone, A., & Garcia-Fontes, W. (2009). *The quality of the Catalan and Spanish education systems: A perspective from PISA* (IESE Business School working paper WP-810). Retrieved from http://ssrn.com/abstract=1513214

Comber, L. C., & Keeves, J. P. (1973). *Science education in nineteen countries: An empirical study*. Stockholm: Almqvist & Wiksell.

Converse, J. M., & Presser, S. (1986). *Survey questions: Handcrafting the standardized questionnaire*. Beverly Hills, CA: Sage.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method* (4th ed.). Hoboken, NJ: Wiley.

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher, 44*, 237–251.

Duncan, G. J., & Murnane, R. J. (Eds.). (2011). *Whither opportunity? Rising inequality, school, and children's life chances*. New York: Russell Sage.

Dweck, C. S., Walton, G. M., & Cohen, G. L. (2011). *Academic tenacity: Mindsets and skills that promote long–term learning*. White paper prepared for the Gates Foundation, Seattle, WA.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75–146). San Francisco: Freeman.

Eccles, J. S., Wigfield, A., Midgley, C., Reuman, D., MacIver, D., & Feldlaufer, H. (1993). Negative effects of traditional middle schools on students' motivation. *Elementary School Journal, 93*, 553–574.

Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., et al. (2012). *Teaching adolescents to become learners. The role of noncognitive factors in shaping school performance: A critical literature review*. Chicago: University of Chicago Consortium on Chicago School Research.

Feeney, J. R., & Goffin, R. D. (2015). The overclaiming questionnaire: A good way to measure faking? *Personality and Individual Differences, 82*, 248–252.

Forster, M. (2004). *Measuring the social outcomes of schooling: What does ACER research tell?* Retrieved from http://research.acer.edu.au/research_conference_2004/4/

Fowler, F. J. (2009). *Survey research methods* (4th ed.). Thousand Oaks, CA: Sage.

Freedman, D. A. (1999). Ecological inference and the ecological fallacy. *International Encyclopedia of the Social & Behavioral Sciences, 6*, 4027–4030.

Gabrieli, C. (2015, March). Non-cognitive skills and education policy: Research and practice considerations. In: M. R. West & G. J. Whitehurst (Chairs), *Ready to be counted? Incorporating noncognitive skills into education policy.* Symposium at the Brown Center on Education Policy, Washington, DC.

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly, 73*(2), 349–360.

Gehlbach, H. (2015a, April 15). Name that baby: Why 'non-cognitive' factors need a new name. *Education Week Blog.* Retrieved from http://blogs.edweek.org/edweek/rick_hess_straight_up/2015/04/non-cognitive_factors_need_new_name.html

Gehlbach, H. (2015b). Seven survey sins. *The Journal of Early Adolescence*. doi:10.1177/0272431615578276.

Ginsburg, A., & Smith, M. S. (2013, December). *Key education indicators (KEI): Making sense of NAEP contextual variables*. Presentation at the National Assessment Governing Board Meeting, Baltimore.

Gorey, K. M. (2001). Early childhood education: A meta-analytic affirmation of the short-and long-term benefits of educational opportunity. *School Psychology Quarterly, 16*(1), 9.

Heckman, J. J., & Kautz, T. (2013). *Fostering and measuring skills: Interventions that improve character and cognition* (IZA discussion paper no. 7750). Bonn, Germany: Institute for the Study of Labor.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24*, 411–482.

Heine, S. J., Kitayama, S., Lehman, D. R., Takata, T., Ide, E., Leung, C., et al. (2001). Divergent consequences of success and failure in Japan and North America: An investigation of self-improving motivations and malleable selves. *Journal of Personality and Social Psychology, 81*(4), 599–615.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Charlotte, NC: Information Age.

Hopkins, D., & King, G. (2010). Improving anchoring vignettes: Designing surveys to correct interpersonal incomparability. *Public Opinion Quarterly, 74*(2), 201–222.

Jakubowski, M., Patrinos, H. A., Porta, E. E., & Wisniewski, J. (2010). *The impact of the 1999 education reform in Poland* (Policy research working paper 5263). Human Development Network Education. Washington, DC: The World Bank.

Jencks, C. (1979). *Who gets ahead? The determinants of economic success in America*. New York: Basic Books.

Kamenetz, A. (2015, May 28). *Nonacademic skills are key to success. But what should we call them?* nprEd. Retrieved from http://npr.org/sections/ed/2015/05/28/4046847/non-academic-skills-are-key-to-success-but-what-should-we-call-them

Kaplan, D., & Wu, D. (2014, April). Imputation issues relevant to context questionnaire rotation. In J. P. Bertling (Chair), *Spiraling contextual questionnaires in educational large-scale assessments*. Symposium conducted at the meeting of the National Council on Measurement in Education, Philadelphia.

Kapteyn, A., Smith, J. P., & Van Soest, A. (2007). Vignettes and self-reports of work disability in the US and the Netherlands. *American Economic Review, 97*, 461–473.

Kautz, T., Heckman, J. J., Diris, R., ter Weel, B., & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success*. Cambridge, MA: National Bureau of Economic Research.

Keehner, M., Agard, C., Berger, M., Shu, Z., Bertling, J., & Carney, L. (2014). *Analyzing data from the NAEP TEL Wells task: Potential reporting insights from interactions, context, and scores* (Research memorandum on NAEP TEL Task Component). Princeton, NJ: Educational Testing Service.

King, G., & Wand, J. (2007). Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes. *Political Analysis, 15*, 46–66.

Kristensen, N., & Johansson, E. (2008). New evidence on cross-country differences in job satisfaction using anchoring vignettes. *Labour Economics, 15*, 96–117.

Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction, 17*(5), 494–509.

Kyllonen, P. C., & Bertling, J. P. (2013). Innovative questionnaire assessment methods to increase cross-country comparability. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.),

Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis (pp. 277–286). Boca Raton, FL: CRC Press.

Kyllonen, P. C., & Bertling, J. P. (2015, April). *Advances in measuring 21st century skills: Constructs, development, and scoring.* Training session presented at the meeting of the National Council on Measurement in Education, Chicago.

Ladd, G. W., Kochenderfer-Ladd, B. K., Visconti, K. J., & Ettekal, I. (2012). Classroom peer relations and children's social and scholastic development: Risk factors and resources. In A. M. Ryan & G. W. Ladd (Eds.), *Peer relationships and adjustment at school* (pp. 11–49). Charlotte, NC: Information Age Press.

Lareau, A. (2011). *Unequal childhoods: Class, race, and family life* (2nd ed.). Berkeley, CA: University of California Press.

Lareau, A., & Weininger, E. B. (2003). Cultural capital in educational research: A critical assessment. *Theory and Society, 32*, 567–606.

Lent, R. W., & Brown, S. D. (2006). On conceptualizing and assessing social cognitive constructs in career research: A measurement guide. *Journal of Career Assessment, 14*(1), 12–35.

Levin, H. M. (2012). More than just test scores. *Prospects, 42*(3), 269–284.

Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction, 19*(6), 527–537.

Maaz, K., Baumert, J., Neumann, M., Becker, M., & Dumont, H. (Eds.). (2013). *Die Berliner Schulstrukturreform: Bewertung duch die beteiligten Akteure und Konsequenzen des neuen Uebergangsverfahrens von der Grundschule in die weiterfuehrenden Schulen*. Muenchen: Waxmann.

Magnuson, K. A., & Votruba-Drzal, E. (2009). Enduring influences of childhood poverty. In S. Danziger & M. Cancian (Eds.), *Changing poverty changing policies* (pp. 153–179). New York: Russell Sage.

Maitland, A., Sun, H., Tourangeau, R., Almonte, D. E., & Bertling, J. P.(2015, March). *Exploration of matrix questions on fourth-grade students using eye-tracking*. Presentation at the NAEP Questionnaire Standing Committee Meeting, Washington, DC.

McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. (2007). Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology, 60*(1), 63–91.

McDonnell, L. M. (1995). Opportunity to learn as a research concept and a policy instrument. *Educational Evaluation and Policy Analysis, 17*(3), 305–322.

Mickelson, R. A. (1990). The attitude-achievement paradox among black adolescents. *Sociology of Education, 63*(1), 44–61.

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research, 3*, 112–131.

Minor, E. C., Desimone, L. M., Spencer, K., & Phillips, K. J. (2015). A new look at the opportunity-to-learn gap across race and income. *American Journal of Education, 121*(2), 241–269.

Monseur, C., & Bertling, J. P. (2014, April). Questionnaire rotation in international surveys: Findings from PISA. In J. P. Bertling (Chair), *Spiraling contextual questionnaires in educational large-scale assessments*. Symposium conducted at the meeting of the National Council on Measurement in Education, Philadelphia.

Monseur, C., & Lafontaine, D. (2012). Structure des systèmes éducatifs et équité : un éclairage international. In M. Crahay (Ed.), *Pour une école juste et efficace* (pp. 185–219). Brussels, Belgium: De Broeck.

Mõttus, R., Allik, J., Realo, A., Pullmann, H., Rossier, J., Zecca, G., et al. (2012). The effect of response style on self-reported conscientiousness across 20 countries. *Personality and Social Psychology Bulletin, 38*, 1423–1436.

Naemi, B., Gonzalez, E., Bertling, J. P., Betancourt, A., Burrus, J., Kyllonen, P. C., et al. (2013). Large-scale group score assessments: Past, present, and future. In D. Saklofske & V. Schwean (Eds.), *Oxford handbook of psychological assessment of children and adolescents*. Cambridge, MA: Oxford University Press.

National Assessment Governing Board. (2012). *Policy statement on NAEP background questions and the use of contextual data in NAEP reporting*. Washington, DC: U.S. Department of Education, National Assessment Governing Board.
National Council of Teachers of Mathematics. Commission on Standards for School Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: NCTM.
Nyhus, E. K., & Pons, E. (2005). The effects of personality on earnings. *Journal of Economic Psychology, 26*(3), 363–384.
O'Connor, M. C., & Paunonen, S. V. (2007). Big five personality predictors of post-secondary academic performance. *Personality and Individual Differences, 43*(5), 971–990.
OECD. (2010). *Pathways to success: How knowledge and skills at age 15 shape future lives in Canada*. Paris: OECD Publishing.
OECD. (2011). *How do education systems and schools select and group students? In PISA 2009 at a Glance*. Paris: OECD Publishing.
OECD. (2012). *Learning beyond fifteen: Ten years after PISA*. Paris: OECD Publishing.
OECD. (2013a). *Education at a glance 2013: OECD indicators*. Paris: OECD Publishing.
OECD. (2013b). *PISA 2012 Assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. Paris: OECD Publishing.
OECD. (2013c). *PISA 2012 results: Excellence through equity: Giving every student the chance to succeed (Volume II)*. Paris: OECD Publishing.
OECD. (2013d). *PISA 2012 results: Ready to learn: Students' engagement, drive and self-beliefs (Volume III)*. Paris: OECD Publishing.
OECD. (2013e). *PISA 2012 results: What makes schools successful? Resources, policies and practices (Volume IV)*. Paris: OECD Publishing.
OECD. (2013f). *PISA 2015: Draft questionnaire framework*. Paris: OECD Publishing. Retrieved from http://www.oecd.org/pisa/pisaproducts/PISA-2015-draft-questionnaire-framework.pdf.
OECD. (2014a). *Education at a glance 2014: OECD indicators*. Paris: OECD Publishing.
OECD. (2014b). *PISA 2012 technical report*. Paris: OECD Publishing.
Oostrom, J. K., De Soete, B., & Lievens, F. (2015). Situational judgment testing. In I. Nikolaou & J. K. Oostrom (Eds.), *Employee recruitment, selection, and assessment: Contemporary issues for theory and practice* (pp. 172–189). New York: Psychology Press.
Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology, 34*(1), 29–48.
Paulhus, D. L., & Dubois, P. J. (2014). Application of the overclaiming technique to scholastic assessment. *Educational and Psychological Measurement, 74*, 975–990.
Paulhus, D. L., Harms, P. D., Bruce, M. N., & Lysy, D. C. (2003). The over-claiming technique: Measuring self-enhancement independent of ability. *Journal of Personality and Social Psychology, 84*(4), 890–904.
Paunonen, S. V., & Ashton, M. C. (2001). Big five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology, 81*(3), 524–539.
Pekrun, R., vom Hofe, R., Blum, W., Frenzel, A. C., Götz, T., & Wartha, S. (2007). Development of mathematical competencies in adolescence: The PALMA longitudinal study. In M. Prenzel (Ed.), *Studies on the educational quality of schools: The final report on the DFG priority programme* (pp. 17–37). Münster, Germany: Waxmann.
Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 63*(2), 167–199.
Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322–338.
Prenzel, M. (2012). Empirische Bildungsforschung morgen: Reichen unsere bisherigen Forschungsansätze aus? In M. Gläser-Zikuda, T. Seidel, C. Rohlfs, A. Gröschner, & S. Ziegelbauer (Eds.), *Mixed Methods in der empirischen Bildungsforschung* (pp. 273–285). Münster, Germany: Waxmann.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta–analysis. *Psychological Bulletin, 138*(2), 353–387.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2*, 313–345.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review, 15*, 351–357.

Rychen, D. S., & Salganik, L. H. (Eds.). (2003). *Defining and selecting key competencies*. Cambridge, MA: Hogrefe & Huber.

Salomon, J. A., Tandon, A., & Murray, C. J. L. (2004). Comparability of self-rated health: Cross-sectional multi-country survey using anchoring vignettes. *British Medical Journal, 328*, 258–261.

Santos, D., & Primi, R. (2014). *Social and emotional development and school learning. A measurement proposal in support of public policy*. São Paulo, Brazil: Ayrton Senna Institute.

Schiepe-Tiska, A., Heine, J. -H., Lüdtke, O., Seidel, T., & Prenzel, M. (2015). *Cognitive and noncognitive outcomes in mathematics classrooms and their relationship with instructional quality. Unterrichtswissenschaft.* Manuscript submitted for publication.

Schunk, D. H., & Mullen, C. A. (2013). Motivation. In J. Hattie & E. M. Anderman (Eds.), *International guide to student achievement* (pp. 67–69). New York: Routledge.

Seidel, T., & Reiss, K. (2014). Lerngelegenheiten im Unterricht. In T. Seidel & A. Krapp (Eds.), *Paedagogische Psychologie* (pp. 253–276). Weinheim, Germany: Beltz.

SES Expert Panel. (2012). *Improving the measurement of socioeconomic status for the National Assessment of Educational Progress: A theoretical foundation* (White Paper prepared for the National Center for Education Statistics). Washington, DC: U.S. Department of Education.

Smith, M. S., Chudowsky, N., Ginsburg, A., Hauser, R., Jennings, J., & Lewis, S. (2012). *NAEP background questions: An underused national resource.* Report to the National Assessment Governing Board by the expert panel on strengthening the NAEP background questions. Retrieved from http://www.nagb.org/content/nagb/assets/documents/publications/expert-panel-naep-bq-report.pdf

Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean–level and rank–order stability of the Big Five. *Journal of Personality and Social Psychology, 101*, 862–882.

Stiglitz, J. E., Sen, A., & Fitoussi, J. P. (2010). *Report by the commission on the measurement of economic performance and social progress*. Paris: Commission on the Measurement of Economic Performance and Social Progress.

Stock, J., & Cervone, D. (1990). Proximal goal-setting and self-regulatory processes. *Cognitive Therapy and Research, 14*(5), 483–498.

United Nations Educational Scientific and Cultural Organization. (2014). *Teaching and learning: Achieving quality for all. EFA global monitoring report*. Paris: UNESCO.

van de Gaer, E., Grisay, A., Schulz, W., & Gebhardt, E. (2012). The reference group effect: An explanation of the paradoxical relationship between academic achievement and self-confidence across countries. *Journal of Cross-Cultural Psychology, 43*(8), 1205–1228.

von Davier, M. (2014). Imputing proficiency data under planned missingness in population models. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis*. Boca Raton, FL: CRC Press.

Weekley, J. A., & Ployhart, R. E. (Eds.). (2006). *Situational judgment tests: Theory, measurement, and application*. Mahway, NJ: Erlbaum.

Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review, 19*(3), 188–202.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*, 461–481.

Wise, S. L., & Kong, X. (2005). Response time effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education, 18*, 163–183.

Ziegler, M., Kemper, C., & Rammstedt, B. (2015). The vocabulary and overclaiming test (VOC-T). *Journal of Individual Differences, 34*(1), 32–40.

Part IV
Conclusions

Chapter 15
Psychosocial Constructs: Knowns, Unknowns, and Where do we go From Here?

Anastasiya A. Lipnevich, Franzis Preckel, and Richard D. Roberts

We must remember that intelligence is not enough. Intelligence plus character—that is the goal of true education. The complete education gives one not only power of concentration, but worthy objectives upon which to concentrate. The broad education will, therefore, transmit to one not only the accumulated knowledge of the race but also the accumulated experience of social living. (King, 1947)

The reader drawn to make a comparison between this opening quote and contemporary educational practice may notice a certain degree of discrepancy: The main focus of post-World War II education has traditionally been on intellectual academic goals rather than skills that could be referred to as "character." The humble intention of this book was to make a small step toward redressing this imbalance by bringing together renowned experts to review the emerging literature on the role, importance, and place of psychosocial skills in K-12 research, public policy, and educational practice. The collection of chapters comprising this volume covers a plethora of topics ranging from theoretical background, assessment, psychometrics, and human development to specific examples of applications of psychosocial skills in educational settings that can be brought to scale. In this concluding commentary, we will summarize the main themes discussed by our contributors and offer ideas

A.A. Lipnevich (✉)
Educational Psychology, Queens College and the Graduate Center, The City University of New York, 365 5th Avenue, New York, NY 10016, USA
e-mail: a.lipnevich@gmail.com

F. Preckel
Department of Psychology, University of Trier, Universitaetsring 15, 54286 Trier, Germany
e-mail: preckel@uni-trier.de

R.D. Roberts
Center for Innovative Assessments, Professional Examination Service, 475 Riverside Drive, Suite 600, New York, NY 10115, USA
e-mail: RRoberts@proexam.org

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8_15

for future research and policy implications, in the process hopefully getting closer toward the "truly complete broad" education that Dr. King envisaged nearly seven decades ago.

We begin by describing psychosocial skills both as predictors of academic outcomes and indicators of a broader definition of academic success, raising issues that we feel may need attention in future scientific discourse. We will also suggest an organizing framework for psychosocial skills that may help to condense the vast number of characteristics into a manageable set of categories. Next, we discuss specific interventions and issues that concern the malleability of psychosocial skills. Finally, we conclude this chapter—and the volume—with a set of recommendations for future educational research, policy, and practice.

15.1 Psychosocial Skills as Effective Predictors of Academic Success

It is no longer disputed that psychosocial skills explain and predict important academic, workplace, and life outcomes. The relative predictive power of these constructs is often discussed in conjunction with that of cognitive skills. Although cognitive skills do predict academic success (e.g., Hezlett et al., 2001; Kobrin, Patterson, Shaw, Mattern, & Barbuti, 2008), job performance (e.g., Schmidt & Hunter, 1998), health (e.g., Deary, Whalley, & Starr, 2003), and marital satisfaction (e.g., Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007), these correlations are consistently far from being perfect (i.e., near unity; correlations range from $r = 0.06$ for mortality to $r = 0.51$ for job performance). Psychosocial skills represent variables that explain additional variance in these and other key outcomes.

Starting from the earliest levels of schooling—preschool—psychosocial skills like conscientiousness predict achievement (e.g., Abe, 2005). This link remains relatively stable in middle school, with self-efficacy, self-concept, and attitudes predicting reading, science, and math grades, as well as scores on both national and international assessments. This relationship persists even after controlling for demographics, school attendance, and educational materials available at home (Campbell, Voelkl, & Donahue, 1997; Lee, Redman, Goodman, & Bauer, 2007). Self-discipline has also been found to predict academic attainment (grades and test scores) beyond cognitive ability for eighth graders (Duckworth & Seligman, 2005). In high school and beyond, meta-analyses have shown that psychosocial factors increment over test scores in predicting GPA, retention, absenteeism, and disciplinary infractions (e.g., Poropat, 2009; Robbins, Lauver, Davis, Langley, & Carlstom, 2004). The key factors include such examples as conscientiousness (e.g., Noftle & Robins, 2007; O'Connor & Paunonen, 2007; Wagerman & Funder, 2006), learning strategies (e.g., Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013), emotional control (e.g., Chamorro-Premuzic & Furnham, 2003; Robbins, Allen, Casillas, Peterson, & Le, 2006), and study habits, skills, and attitudes (e.g., Crede & Kuncel, 2008).

Table 15.1 Relationship of selected psychosocial constructs with academic outcomes

Construct	Outcomes	Chapter and authors
Need for cognition	Reasoning High school GPA College GPA Attendance Test scores (ACT) Domain-specific knowledge Course grades Domain knowledge	Chapter 5 (Jebb, Saef, Parrigon, & Woo)
Creativity	GPA Teacher ratings of performance Domain-specific knowledge Science understanding	Chapter 6 (Kaufman, Beghetto, & Dilley)
Conscientiousness	High school GPA College GPA Attendance Course Grades Tardiness Subject knowledge	Chapter 7 (Kim, Poropat, & MacCann)
Academic self-concept	School grades Test scores GPA	Chapter 8 (Trautwein & Möller)
Self-regulated learning	GPA Course grades Test scores Teacher ratings of performance	Chapter 9 (Bembenutti, White, & DiBenedetto)
Motivation	High school GPA College GPA Test scores (SAT, ACT) First-year retention in college	Chapter 10 (Hulleman, Barron, Kosovich, & Lazowski)
Academic emotions	GPA Subject-specific course performance (math, science, languages)	Chapter 11 (Goetz & Bieg)
Resilience	GPA Attendance Dropout rates	Chapter 12 (Prince-Embury, Keefer, & Saklofske)

Psychosocial variables described in this volume are numerous, and each of them has been shown to relate to key academic outcomes. In some meaningful sense then, the authors contributing to this volume have selected and extended the constructs discussed in the aforementioned passages. Table 15.1 lists the constructs described in detail throughout the current volume and the academic outcomes to which they most strongly relate. Table 15.2 is a companion piece, where we systematically list promising measures and approaches to assessment of each of these constructs, and Table 15.3 summarizes interventions and suggestions for skill development discussed in the current volume.

The relationship of psychosocial skills with meaningful outcomes continues on well beyond schooling, extending into the workforce and throughout an individuals'

Table 15.2 Examples of assessments of psychosocial constructs discussed in the volume and elsewhere

Psychosocial skill	Assessments
Empathy (Chapter 3)	Bryant Index of Empathy Measurement for Children and Adolescents, Southampton Test of Empathy for Preschoolers (STEP), Basic Empathy Scale (BES), Interpersonal Reactivity Index (IRI), Multifaceted Empathy Test (MET), Balanced Emotional Empathy Scale (BEES), skin conductance, respiration, measure of daily helping
Self-esteem (Chapter 3)	Rosenberg's Self-Esteem Scale, Coopersmith Self-Esteem Inventory, Piers-Harris Children's Self-Concept Scale (CSCS), Personal Evaluation Inventory (PEI)
Self-directed learning (Chapter 3)	Guglielmino's Self-Directed Learning Readiness Scale (SDLRS), Self-Directed Learning Scale (SDLS), Self-Directed Learning with Technology Scale (SDLTS)
Leadership (Chapter 3)	Multifaceted Leadership Questionnaire (MLQ), Roets Rating Scale for Leadership (RRSL), Scales for Rating the Behavioral Characteristics of Superior Students (SRBCSS), Vanderbilt Assessment for Leadership in Education
Civic knowledge (Chapter 3)	United States Citizenship and Immigration Services (USCIS) Naturalization Test, National Assessment of Educational Progress (NAEP), International Association for the Evaluation of Educational Achievement Civic Education Study
Teamwork and cooperation (Chapter 3)	Teamwork Knowledge Skill and Ability Test (TKSA), Teamwork Competency Test (TCT), Individual Performance in Teams Scale (IPIT), Situational Judgment Tests (SJT), self-report and teacher-rating scale
Ethics (Chapter 3)	Defining Issues Test, Schwartz Value Scale (SVS), Ethical Priority Test, Ethical Sensitivity Scale Questionnaire, Aristotelian Ethical Behavior in Leisure Scale (AEBLS)
Emotional intelligence (Chapter 3, 11, 13)	Bar-On EQ-i, EQ-i youth form, Situational Test of Emotional Management for Youth (STEM-Y), Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT), MSCEIT-YV, Children's Worry and Anger Management Scales (CWMS, CAMS), Children's Sadness Management Scale (CSMS), Trait Emotional Intelligence Questionnaire (TEIQue; Child Form, TEIQue-CF; Adolescent Short Form, TEIQue-ASF), "Guess Who? Peer Assessment Technique"
Student motivation (Chapter 3, 10)	Motivations for Reading Questionnaire (MRQ), Expectancy-Value-Cost (EVC) Scale, Intrinsic Motivation Inventory (IMI), PISA study, Preschool Reading Attitude Scale (PRAS), Emergent Readers Motivation and Reading Scale (ERMAS), Student Motivation and Engagement Scale—High School (MES-HS)
Need for cognition (Chapter 5)	The Need for Cognition Scale (NCS), Need for Cognition in Early Adolescence (German adaptation)
Creativity (Chapter 6)	Self-reported creativity (e.g., creative self-efficacy), creativity activity checklists (Kaufman Domains of Creativity Scale, K-DOCS), teacher ratings (Scales for Rating the Behavioral Characteristics of Superior Students), divergent thinking tests (e.g., Torrance Test of Creative Thinking, TTCT), creative problem-solving tasks (Remote Associates Test)
Conscientiousness (Chapter 7)	NEO Personality Inventory—Revised (NEO-PI-R), Big Five Aspect Scale, Sixteen Personality Factor Questionnaire, Abridged Big Five Dimensional Circumplex Scale, California Personality Inventory, Hogan Personality Inventory, Jackson Personality Inventory—Revised, Multidimensional Personality Questionnaire, HEXACO Personality Inventory, Conceptual Synthesis of Conscientiousness Facets Across Models

(continued)

Table 15.2 (continued)

Psychosocial skill	Assessments
Grit (Chapter 7)	Short Grit Scale (Grit-S)
Academic self-concept (Chapter 8)	Academic Self-Description Questionnaire (ASDQ), Piers-Harris Children's Self-Concept Scale (CSCS), Joseph Pre-School and Primary Self-Concept Screening Test (JPPSST), Pictorial Scale of Perceived Competence and Social Acceptance (PSPCSA)
Self-regulated learning (Chapter 9)	Motivated Strategies for Learning Questionnaire (MSLQ), Self-Regulated Learning Interview Schedule (SRLIS), observations of overt behavior, interview evidence, think-aloud protocols, traces of mental events and processes, situational manipulations, recording student motivation strategies as they work, keeping diaries, microanalytic assessments
Academic emotions (Chapter 11)	Ekman's "Facial Action Coding System" (FACS), Positive and Negative Affect Schedule (PANAS), Academic Emotions Questionnaire (AEQ), Self-Assessment Manikin (SAM), State-Trait Anxiety Inventory (STAI), physiological measures (skin conductance, heart rate, cortisol), imaging techniques (fMRI, EEG), FaceReader
Resiliency (Chapter 12)	Resiliency Scales for Children and Adolescents (RSCA): a profile of personal strengths, ClassMaps Survey

Table 15.3 Examples of interventions and recommendations for development of psychosocial constructs discussed in the volume

Psychosocial skill	Interventions/recommendations
Emotional intelligence (Chapters 3, 11, 13)	Promoting Alternate Thinking Strategies (PATHS) program and Promotion of EI in Learning and Achievement Situations (PEILAS) model build knowledge of emotions and strategies for emotion regulation; RULER (universal SEL approach that integrates EI into core academic curriculum; skills include recognizing, understanding, labeling, expressing, and regulating emotions)
Student motivation (Chapters 3, 10)	Attribution retraining, growth mindset interventions (target perceptions about student capacity to learn), opportunities for choice in the classroom (e.g., homework assignments), value affirmation, interventions targeting students' sense of belonging in academic environment, The Wheel (13-module motivation and engagement framework), Concept-Oriented Reading Instruction (CORI)
Need for cognition (Chapter 5)	Intellectual discussions with faculty/peers during critical periods, increase student's sense of self-efficacy, programs that increase depth and frequency of student's reading
Creativity (Chapter 6)	Incorporate mini-c learning experiences, supportive and balanced task-specific feedback, explicitly teaching creative thinking strategies, foster creative metacognition in students
Conscientiousness (Chapter 7)	Enhancement of self-regulation (Tools of the Mind program), attentional bias modification, mindfulness training, teacher conscientiousness
Grit (Chapter 7)	Present tasks as investments in long-term goals, similar interventions to conscientiousness
Academic self-concept (Chapter 8)	Comprehensive school systems and teachers that bolster self-concept, temporal comparisons instead of social comparisons in giving feedback, positive motivational cues, supportive attributions for success and failure, safe and trusting student-teacher relationships

(continued)

Table 15.3 (continued)

Psychosocial skill	Interventions/recommendations
Self-regulated learning (Chapter 9)	Foster a learning environment where students focus on personal progress and view errors as learning opportunities, modeling desired behaviors, teaching goal setting (K-2), modeling and integrating help-seeking strategies (3–5), modeling self-assessment (6–8), encouraging students to evaluate their own work and self-monitor progress (9–12), and using computer-based learning environments
Academic emotions (Chapter 11)	Teach learners that academic emotions are both controllable and valuable, teach learners about emotions and strategies for regulating emotions, define and classify academic emotions, build emotion vocabulary, increase knowledge about the effects of academic emotions, and focus on mastering learning material and working through challenges
Resiliency (Chapter 12)	Addressing emotional reactivity, social and emotional learning (SEL) programs that target self-management, social awareness, relationships skills, and responsible decision-making; give students opportunities to practice social and emotional skills inside and outside of the classroom (also at home); Responsive Classroom approach; ClassMaps; cultivating "islands of competence"/mastery experiences; effort attributions; positive feedback that focus on strengths

life. Psychosocial variables demonstrate significant relationships with a vast array of variables, important for individuals' development and functioning (see Hough & Oswald, 2008; Matthews, Zeidner, & Roberts, 2006; Ozer & Benet-Martinez, 2006; Roberts et al., 2007 for reviews). These include happiness (e.g., Diener & Lucas, 1999), health (e.g., Bogg & Roberts, 2004), longevity (e.g., Roberts et al., 2007), job performance (e.g., Barrick, Mount, & Judge, 2001), job satisfaction (e.g., Judge, Heller, & Mount, 2002), labor economic outcomes (e.g., wages, employment, incarceration rates; see e.g., Heckman & Rubinstein, 2001; Heckman, Malofeeva, Pinto, & Savelyev, 2007), marital satisfaction (e.g., Watson, Hubbard, & Weise, 2000), peer relationships (e.g., Jensen-Campbell et al., 2002), as well as behavioral problems (e.g., Ge & Conger, 1999) and psychological disorders (e.g., Trull & Sher, 1994). All in all, evidence from empirical research supports the claim that psychosocial skills are important predictors of academic, workforce, and—ultimately—life success.

15.2 Psychosocial Skills as Meaningful Outcomes in Their Own Right

The goal of finding effective ways to predict students' academic success has been high on the priority list of numerous researchers in the field of education. Scholars have attempted to discover additional predictors of academic achievement that should be considered during the admissions process or when evaluating successes on individual, school, or programmatic levels (e.g., Burton & Ramist, 2001;

Harackiewicz, Barron, Tauer, & Elliot, 2002). Discovery and definition of new predictors of students' academic success, however, is a futile exercise, unless there is a comprehensive formulation of *what* constitutes such success. Until recently, the plethora of newly proposed and traditional predictor variables was used in conjunction with a very narrow set of academic outcomes. Specifically, the criteria for success that have been most popular among the researchers were first-year and cumulative college GPA, course grades, college or high school graduation, and attendance rates (see, e.g., Camara & Echternacht, 2000; Camara & Kimmel, 2005; Harackiewicz et al., 2002). Such a mismatch between the predictor and the outcome space stalls potential developments in the field and makes the impact of inquiries theoretically and practically less meaningful.

Throughout this volume, the contributing authors conveyed the message that psychosocial skills are valuable in their own right. That is, teaching students time management, self-regulation, motivation, creativity, and other psychosocial characteristics should be an explicit goal of education. In fact, as Stemler and DePascale note in Chapter 3 of this volume, these skills have already become an ingrained part of the core mission of many educational establishments. Further, as Bertling, Borgonovi, and Almonte (Chap. 14) show, these skills are becoming the focus of many national and international large-scale group score assessments and are regarded as key outcomes of education. Hence, there is a growing recognition that psychosocial skills should be at the very core of instructional programs irrespective of their links to GPA or other "traditional" academic outcomes. After all, it is virtually impossible to build a cogent argument that would somehow disprove the pivotal role of academic self-concept, emotional control, creativity, or motivation, to name a few, for individuals' success. So let us look at psychosocial skills as outcomes of educational programs and let us review approaches to defining the criterion space in education that includes such skills.

Many prestigious colleges and universities include psychosocial skills into their definitions of success. Such definitions range from broad goals such as helping develop the whole person and preparing students for the adult world to specific lists of skills and attributes they expect of their graduates. Examples of the latter include students' ability to generate original ideas and solutions, function in an intercultural context, and respect and value individual differences (see, e.g., Acrum, 1992; Rigol, 2003). Many institutions now claim that to the extent possible, admissions decisions should be validated on this broader set of criteria. Thus, the chapters that focus on specific psychosocial constructs discussed in this volume are often consistent with a reconceptualization of psychosocial skills as explicit goals of education.

An interesting approach to identifying a comprehensive list of outcomes of educational programs was undertaken by Oswald, Schmitt, Kim, Ramsay, and Gillespie (2004). The researchers examined educational objectives and mission statements from 35 colleges and universities in the United States in search for common themes that institutions stipulated. The authors condensed all the variables into twelve relevant criteria of college success. The identified dimensions were:

1. Knowledge, learning, and mastery of general principles
2. Continuous learning, intellectual interest, and curiosity

3. Artistic and cultural appreciation and curiosity
4. Multicultural tolerance and appreciation (showing tolerance and openness)
5. Leadership (demonstrating skills in a group)
6. Interpersonal skills (communicating and dealing well with others)
7. Social responsibility, citizenship, and involvement
8. Physical and psychological health (avoiding unhealthy behaviors, having a set of effective coping mechanisms for dealing with stress)
9. Career orientation (establishing, prioritizing, and working toward goals)
10. Adaptability and life skills (adapting well to changes, dealing with problems)
11. Perseverance (committing to goals, regardless of their difficulty)
12. Ethics and integrity (having a well-developed set of values)

If mission statements are to be regarded as formal documents with explicated goals of educational institutions, then the proper identification and further assessment of corresponding criteria for success should be in place. Undoubtedly, there is evidence demonstrating that college success is a multidimensional construct that should not be gauged solely on students' GPA, grades, and retention. The latter statement is bolstered by evidence that theoretically derived components of college success were predicted by distinct indicators and covaried with other variables (Oswald et al., 2004). Stemler and DePascale (Chap. 3) provide a detailed review of the criteria of educational success as defined through mission statements at the K-12 level and beyond and offer a list of assessments that can be used to gauge them. It is important to note differences and commonalities between tertiary education, on the one hand, and primary and secondary, on the other. The main differences may stem from the selection procedures: Universities and colleges, in general, are more selective than primary, middle, and high schools (private schools, notwithstanding). Further, university and college students have a longer learning history and more experience with formal education and have accumulated wider academic knowledge and skills as compared to students of the K-12 system. Nevertheless, there are also commonalities between primary, secondary, and tertiary education. First, all three levels provide formal instruction with the aim of preparing student for the later (work) life. Second, the biological and neurological basis of learning (i.e., basic brain structures and functions) does not change qualitatively from secondary to higher education.

The importance of psychosocial skills is also recognized outside of the academic milieu. In fact, researchers in education draw upon current theories of job performance to reevaluate the domain of students' performance in schools and colleges. The organizational psychology literature presents several models of competencies that describe criteria for success in the workplace (see, e.g., Robertson, Callinan, & Bartram, 2002). In a report published by The Conference Board, Partnership for 21st Century Skills, Society for Human Resource Management, and Corporate Voices for Working Families, *Are They Really Ready to Work?* researchers identified the skills recognized as most important and/or that will be increasing in importance over the next decades as critical to workplace success based on responses from 400+ surveys and twelve interviews with HR professionals and executives in

the business community (Casner-Lotto & Barrington, 2006). Psychosocial skills such as work ethic, teamwork, oral communication, leadership, and creativity were listed as some of the most important personal skills for success in the workplace in the twenty-first century. These skills are reviewed in the current volume with our contributors offering specific suggestions for psychosocial skill development and assessment in K-12 education. Interestingly, each of these skills was rated as more important, compared to the skills traditionally taught and assessed in school, such as writing, reading comprehension, English, and math.

To conclude our discussion of models of success that encompass psychosocial skills, we will review the final set of characteristics that may help us to define success in education. Bartram, Robertson, and Callinan (2002) (see also Kurz & Bartram, 2002) put forward a list of eight criteria, or competency factors, aptly entitled the Great Eight. These competencies were derived by employing factor analysis and multidimensional scaling analysis to categorize supervisor, self-assessment, and overall job performance ratings. As Bartram et al. (2002) note, this approach represented a criterion-centered model, rather than predictor-centered one, when the data obtained through cognitive measures, motivation, and personality questionnaires were analyzed. These factors capture a broad range of skills and attributes and include:

1. Leading and deciding (providing leadership, initiating action)
2. Supporting and cooperating (team working, supporting)
3. Interacting and presenting (relating, communicating, influencing)
4. Analyzing and interpreting (problem solving, writing, applying expertise and technology)
5. Creating and conceptualizing (learning and researching)
6. Organizing and executing (planning, delivering quality, persevering)
7. Adapting and coping (coping with stress)
8. Enterprising and performing (achieving results)

The Great Eight factors have been rigorously tested, and the structure has been replicated in a number of different data sets. Bartram et al. (2002) present evidence from 33 validation studies and demonstrate support of the eight-component structure. Ability and personality variables predict the Great Eight factors differently and in a meaningful fashion. The researchers conclude that the Great Eight model provides a useful framework for testing predictor-criterion contingencies with both personality and cognitive indicators used as predictors.

The Great Eight model appears to be a good solution to the criterion problem in the workplace. Although some attempts have been made by educational researchers to construct similar models that would be as structurally stable and theoretically meaningful, in general these have neither generalized nor been widely adopted. It is possible that defining the exact set of skills that are critical for academic success at all levels is an unattainable endeavor, but it is something that researchers should attempt to achieve in future programmatic endeavors. In concluding this section, we contend that the skills listed throughout this volume should certainly be considered in any such model and become key outcomes of many educational programs.

15.3 An Organizing Framework for Psychosocial Skill Assessment and Development

This section comments on an organizing framework for the key psychosocial factors discussed as early as Chapter 1. Burrus and Brenneman (Chap. 1) attempted to organize the collection of psychosocial skills into a concise model. The authors discuss the Big Five as the key organizing framework for psychosocial skills. The authors propose to further divide the Big Five into three categories, critical for K-12 students: performance skills, interpersonal skills, and self-management skills. The performance skills ("getting along with school") are generally aligned with conscientiousness and openness to experience. According to Burrus and Brenneman (Chap. 1), skills that fall under the umbrella of the performance skills include grit (Chap. 7), creativity (Chap. 6), curiosity (Chap. 5), time management (Chap. 7), and goal setting (Chap. 9 and 10). The second category proposed by Burrus and Brenneman (Chap. 1) is self-management skills ("getting along with yourself"). Self-management skills are all related to the emotional stability factor of the Big Five and include skills like self-efficacy (Chap. 9 and 10; see also Chap. 8 on academic self-concept), test anxiety (Chap. 11), and coping (Chap. 11 and 12). Burrus and Brenneman (Chap. 1) also suggest that interpersonal skills ("getting along with others") can be indexed by such skills as leadership (related to conscientiousness and openness to experience) and teamwork (related to agreeableness and emotional stability). The authors further propose that cross-cultural competence (Chap. 1) and emotional intelligence (Chap. 11 and 13) fall under the interpersonal skills category.

At this juncture, it would be disingenuous not to discuss other alternative frameworks for synthesizing key psychosocial skills for educational research, policy, and practice. For example, the Partnership for 21st Century Skills considers the four Cs—critical thinking, communication, collaboration, and creativity—as core, while the Collaborative for Academic, Social, and Emotional Learning (CASEL) highlights five "competency clusters": self-awareness, self-management, social awareness, relationship skills, and responsible decision-making. Tony Wagner's (2010) bestseller, *The Global Achievement Gap*, highlights seven "survival skills": problem solving and critical thinking, collaboration across networks and leading by influence, agility and adaptability, initiative and entrepreneurship, effective written and oral communication, accessing and analyzing information, and curiosity and imagination. Another bestseller, Paul Tough's (2013) *How Children Succeed* champions grit, curiosity, and the hidden power of character in its subtitle. And in 2012, the National Research Council, in its landmark report *Education for Life and Work*, attempted to cut through the morass by declaring three clusters of competencies: the cognitive, the interpersonal, and the intrapersonal.

Elsewhere, we have made a case that each of these models, however, may also be subsumed under the Big Five factor framework (Roberts, Martin, & Olaru, 2015). Moreover, whereas the Big Five is supported by a large volume of compelling meta-analytic data showing its utility for education, workforce, and life across a wide array of the outcomes, in many countries, other frameworks are based on

isolated studies, have limited empirical support, and/or appear to predict only specific academic outcomes. Indeed, we suggest that any alternative model/framework to be considered should exceed the Big Five with respect to the following criteria:

1. Appropriately, evidence, theory, and policy driven
2. Well-documented, extensive, and programmatic record of validity support, including meta-analytic evidence (extra weighting would come in the form of causal modeling through longitudinal designs; note however, many of these longitudinal studies use the Big Five framework [see Chap. 4] so this may be a heavy lift)
3. Extensive predictive scope (i.e., provide recommendations or evidence tied not just to education, but also workforce and everyday life considerations)
4. Global, cross-cultural relevance (i.e., not limited just to data collected in educational systems in the United States)
5. Ability to generate actionable recommendations and also have intervention programs that are consistent with (1)–(4), such as is documented in many of the chapters of this volume
6. Extent to which a logical case can be made that the alternative model/framework will yield a return on investment greater than the Big Five

Put another way, it is precisely for these reasons that we believe the Big Five (including their facets) serves as a compelling framework for understanding K-12 psychosocial skills, how they would best be measured, and how they might be changed. Unfortunately, unlike cognitive skills, which have a fairly comprehensive taxonomic model underlying it—the Cattell-Horn-Carroll model (see Roberts & Lipnevich, 2011)—with primary mental abilities identified under higher-order constructs, the Big Five facets remain idiosyncratic to particular researchers' preference. We contend that a major undertaking needed in this domain is to clearly document these facets, especially since it appears, that is, at this level successful interventions may be targeted (see, e.g., Kyllonen, Lipnevich, Burrus, & Roberts, 2014). Fortunately, some work has been done to this end (especially for conscientiousness; see, e.g., MacCann, Fogarty, & Roberts, 2012; Roberts, Chernyshenko, Stark, & Goldberg, 2005), but the goal of having a stratum model for psychosocial skills akin to cognitive ability appears some time off in the future.

15.4 On the Issue of Malleability of Psychosocial Skills

Researchers have consistently demonstrated that cognitive ability may not be changed easily (see, e.g., Kyllonen, Roberts, & Stankov, 2008). Conversely, as our contributors note in their respective chapters, most psychosocial skills can be (see Chap. 4, 10, 11, 12, and 13, for review; Table 15.3 lists examples of interventions and recommendations for psychosocial skills development). Walton and Billera (Chap. 4) provide a comprehensive review of literature on the development of personality across the individuals' life. The authors do a tremendous task explicating the complexity of assessing stability and change of psychosocial

characteristics, so we would like to refer the reader to Chapter 4 of this volume for the full account. As it comes to mean-level changes, the authors condense their conclusions to two main points. First, personality continues to change in adulthood (Roberts, Walton, & Viechtbauer, 2006). Studies show significant changes in psychosocial skills during childhood and adolescence (Branje, Frijns, Finkenauer, Engels, & Meeus, 2007; De Fruyt et al., 2006; Klimstra, Hale, Raaijmakers, Branje, & Meeus, 2009; McCrae et al., 2002; Prinzie & Dekovic, 2008; Pullmann, Raudsepp, & Allik, 2006), with young adulthood being the period for the most significant changes. Second, the authors present a rather optimistic picture showing that individuals' personal characteristics generally improve as they age. So, individuals exhibit increased conscientiousness, emotional stability, and social dominance across much of the life course. This supports the so-called maturity principle, which states that people demonstrate an increased capability of being a productive and involved member of society, an increased tendency to be planful and decisive, and a greater propensity for being considerate and charitable (see Caspi, Roberts, & Shiner, 2005). In sum, personal betterment is one more thing that we should look forward to as we age.

The research that Walton and Billera (Chap. 4) review suggests that specific interventions targeting psychosocial skills may be of great use and importance. After all, personality is not set in plaster after the age of five, as James (1890/1981) originally proposed (see, e.g., Terracciano, Costa, & McRae, 2006). Attesting to this point, a recent meta-analysis by Durlak, Weissberg, and Pachan (2010) synthesized results of 75 studies that examined the impact of after-school programs on a number of meaningful outcomes. The authors found that these nonformal learning programs had an overall positive and statistically significant impact on a range of psychosocial characteristics of participating students. The main changes in outcomes took place in three main domains: feelings and attitudes, indicators of behavioral adjustment, and school performance. These authors' also revealed significant increases in children's and adolescents' self-perceptions (e.g., self-concept), bonding to school, positive social behaviors, school grades, and achievement test scores. Further, problem-related behaviors were reduced. In sum, Durlak et al. (2010) conclude that these findings indicate that after-school programs deserve support and recognition and stress the importance of these programs as means to increase a slew of key psychosocial characteristics.

The contributors to this volume did a thorough job of systematically reviewing interventions and approaches to enhancing core psychosocial skills. Thus, Prince-Embury, Keefer, and Saklofske (Chap. 12) provide specific examples of psychosocial interventions aimed at fostering resiliency in the classroom: Responsive Classroom and ClassMaps. Both programs have garnered significant empirical support, and their effectiveness is seemingly beyond debate. Similarly, Torrente, Rivers, and Brackett (Chap. 13) review approaches to improving emotional intelligence in schools. They present a RULER model that focuses on the development of five skills that can be instilled by means of instruction and experience: recognizing emotions (e.g., interpreting nonverbal cues), understanding (e.g., knowing causes and consequences of emotions), labeling (e.g., developing extensive emotion

vocabulary), expressing (e.g., expressing emotion in socially appropriate ways), and regulating emotions (e.g., managing emotional states). Torrente et al. describe a range of specific tools that help to enhance students' emotional intelligence. Goetz and Bieg's (Chap. 11) discussion adds to our understanding of approaches to helping students regulate their emotions. The authors further present a model for the development of emotional intelligence in learning and achievement situations.

Another broad set of psychosocial skills includes investment traits like need for cognition. Jebb, Saef, Parrigon, and Woo (Chap. 5) provide suggestions on how need for cognition, as the tendency to enjoy effortful thinking, may be developed in the classroom. The authors note that high self-efficacy as well as programs focused on the depth and frequency of student reading may support their need for cognition. Kaufman, Beghetto, and Dilley (Chap. 6) review individual and school-level approaches to the development of creativity. From teaching students creative metacognition (Kaufman & Beghetto, 2013) to creating contexts that encourage creative solutions to seemingly mundane problems—the authors provide persuasive evidence suggesting that creativity can be improved.

Academic self-concept is a characteristic that can be enhanced rather economically and without disrupting the routine instructional activities and that has a potential for positive ramifications for the students. In fact, as Trautwein and Moeller (Chap. 8) note, some interventions resulted in self-concepts that increased by half of a standard deviation, which in turn lead to improved academic performance. Due to the fact that high academic self-concepts in one domain can lead to reduced self-concepts in other, contrasting domains, interventions fostering academic self-concept are often included into broader programs that are focused on improving student self-regulation. Bembenutty, White, and DiBenedetto (Chap. 9) describe specific tools for encouraging self-regulated learning in students of various ages. The advice ranges from teaching children at K-3 level how to resist lunch until lunchtime to helping K-6- to K-12-level students examine the consequences of partying before large homework assignments are due.

Motivation is another central psychosocial skill that can be successfully developed through targeted interventions. Hulleman, Kosovich, and Lasowski (Chap. 10) present a meta-analysis of expectancy interventions, value interventions, and cost interventions, as well as multicomponent motivational interventions. The authors discuss an impressive empirical base attesting to the effectiveness of these approaches.

In sum, the corpus of literature on interventions reviewed in this volume suggests relatively unequivocally that psychosocial skills are teachable and learnable and that such interventions can generally be done at scale. This does not suggest, however, that there is still not more work to be done. For one, the practitioner may feel overwhelmed by the large number of programs available, with little attempt to identify redundancies, or provide a clear evaluation of the efficacy of one program relative to another. This issue is clearly outside the scope of this final chapter, but resources exist for making such judgments, such as is provided on the webpage of the Collaborative for Academic, Social, and Emotional Learning (CASEL, see http://www.casel.org/). Secondly, not all contributors made it abundantly clear that the hallmark of effective programs is a systematic process that requires multiple

steps and conditions. Elsewhere we have provided guidelines for the development, implementation, and evaluation of programs of a psychosocial nature (Zeidner, Roberts, & Matthews, 2002). It is perhaps worth briefly reiterating these guidelines here, since while generally met, virtually no intervention program included all elements (certainly as presented in this volume). This is clearly something that needs to occur as these programs move to scale, both domestically and internationally:

1. Base psychosocial skills intervention programs on a solid conceptual framework
2. Carefully specify program goals and behavioral outcomes
3. Fully integrate psychosocial programs into the school educational and instructional curriculum
4. Make provisions for practice and for the transfer of these skills outside the classroom (after-school settings provide an especially important context for transfer)
5. Ensure professional development of program personnel (modeling the targeted behaviors for the student should be a key intervention strategy)
6. Use robust experimental, psychometrically sound designs for assessing program effectiveness

Finally, it is not always clear how these intervention programs can be buttressed effectively by educational policy. Absent in this important impetus, we remain circumspect of these programs' sustainability. In this next section, we attempt to redress this imbalance by integrating all of the preceding sections to provide recommendations for future educational research and program evaluation and how policy might service these endeavors.

15.5 Policy Implications and Future Directions

The contributors of this volume affirm the utmost importance of psychosocial skills for individuals' learning, functioning, and development. Not surprisingly, these skills have been increasingly taken seriously by the educational community, as witnessed by the prominence of standards movements (e.g., Partnership for 21st Century Skills, 2006a, b, see Chap. 3), and their growing role in large-scale international assessments with an attendant impact on policy (see Chap. 14) and even legislation (e.g., Partnership for 21st Century Skills, 2008). In fact, in countries as different as the United States, the United Kingdom, Finland, Korea, Israel, and Singapore, psychosocial skills have been elevated to playing a central role in national curricula. This movement has been fueled, in part, by a new understanding that these skills are critical in the global economy. A number of US states currently mandate that psychosocial skills be part of a standard curriculum, with more states joining in.

To build on the growing momentum, we feel that psychosocial skills are destined to take a more prominent place in school curricula. Psychosocial skills should become explicit in accountability practices and policies, and expectations that educational programs develop and focus on these characteristics should be clear to educators, students, and parents alike. This task, of course, is not simple, but hardly

anything in education is. Changing curriculum, promoting teacher preparation, receiving support from administrators, and educating parents on the extended criteria of their children's success are some of the steps that will have to take place in order to make psychosocial skills an ingrained part of daily instructional practices.

Further, schools in general and teachers in particular should be held accountable for teaching psychosocial skills. Currently, one of the main (and much criticized) criteria of teacher effectiveness is student performance on standardized tests. Adding psychosocial skills to the list of indicators of teacher success may sound intimidating, but it is quite possible that such addition may be welcomed and embraced by educators. After all, students' achievement on a test does not capture the breadth of teachers' contribution to students' learning and development, and broadening the outcome space to include student psychosocial skills may be a boon to all. As Stemler and DePascale note in Chapter 3 of this volume, most teachers are generally in favor of the concept of accountability—as long as such system is aligned with the goals that they deem important.

The next policy implication flows out of the previous one. To make judgments about students' psychosocial skills, such skills should be effectively assessed, and the results of such assessments should be clearly communicated to all the interested parties. Chapter 2 of this volume reviews issues and concerns associated with assessment of psychosocial skills and proposes solutions to solving some of the issues, whereas Chapter 3 offers specific examples of such assessments. Using effective traditional and alternative approaches to gauging psychosocial skills does not have to be an overly daunting task. With the help of technology, in particular, this task may become easily executable, as some program-level precedents show (see Roberts et al., 2015).

As we mentioned earlier in this chapter, after-school programs have been shown to contribute to positive development of psychosocial skills. Music, arts, drama classes, sports, and peer tutoring are some of the many quality after-school experiences that have been shown to enhance a range of psychosocial skills. Making such programs available to as many students as possible may be another avenue for shifting policies.

We may be stating the obvious here, but policy changes should be carefully construed and steeped in solid research foundations. The challenge here is that the domain of psychosocial skills research is very broad as it transcends a number of fields and disciplines that lie on the intersection of psychology, education, economics, and other related disciplines. It is virtually impossible to summarize all research on psychosocial skills, even in the specific context that we chose for this volume—K-12. Despite existing consensus on a number of issues that concern psychosocial skills in K-12, researchers working in the field have enough problems to solve and questions to answer. Let us try to summarize some of the areas in which we think research may most usefully progress and inform policy changes.

Researchers may work on identifying which skills (and when) are the most essential for students. The researchers in this volume describe a number of psychosocial skills that are important for students to cultivate. As Burrus and Brenneman (Chap. 1) suggest, researchers may also want to investigate their relative importance

in a certain developmental phase. In other words, although all of the skills that were mentioned in this book would be nice to have, is there a subset of skills that are absolutely critical to develop? Are there any basic psychosocial skills that set the stage for the development of more complex skills? After all, time and effort are valuable and in many cases quite limited. Hence, having a clear understanding as to which skills should be the focus of attention would be of great theoretical and practical utility. Related to this point, researchers can invest into identifying which psychosocial skills are most malleable. Burrus and Brenneman (Chap. 1) speculate that if, for example, "leaders are born, not made," trying to enhance individuals' leadership skills would be a futile exercise. Teaching students on how to manage their time or how to cope with debilitating anxiety may be more useful.

Another broad area for research concerns the summative as well as the formative assessment of psychosocial skills. Ziegler and Brunner (Chap. 2) and Stemler and DePascale (Chap. 3) discuss a range of important questions that arise when psychosocial skills are to be gauged. For example, researchers should continue their efforts and develop new assessments of these skills that are resistant to faking (see Ziegler, MacCann, & Roberts, 2011 for an extensive review). Some promising options include the use of forced-choice methods (e.g., Stark, Chernyshenko, Drasgow, & White, 2012), anchoring vignettes (Bertling et al., Chap. 14), situational judgment tests (Lipnevich, MacCann, & Roberts, 2013), and assessments embedded in video games (e.g., Shute, Ventura, Kim, & Wang, 2014). Moreover, instituting a system of formative assessment with specific tools and recommendations that would be helpful in developing psychosocial skills in students would be yet another useful topic for research. In a similar way that educators learn to provide feedback on student writing, teachers should be encouraged to provide feedback on psychosocial skills. Both the form and the content of such feedback are a critical topic for investigation.

Finally, choosing a parsimonious taxonomy of psychosocial skills would move the field forward. As we and our contributors noted earlier in this volume (see, e.g., Chap. 1 and 4), the Big Five personality model may serve as an organizing framework of psychosocial skills. However, this model needs to be more fully articulated at the facet level to provide the fine-grained organizing taxonomy of psychosocial skills that might serve researchers, practitioners, and policy makers best. Such a taxonomy would help to expand the criterion space of educational success by defining key characteristics. Researchers could use such framework to integrate scattered findings or to identify blind spots of educational research (e.g., systematic intervention studies for some of these skills). And, as Stemler and DePascale (Chap. 3) rightfully note, educators can be held accountable for developing psychosocial skills, in the same way they are held accountable for developing reading, writing, and arithmetic in their students. Developing such a taxonomy is an ambitious endeavor that presupposes the cooperation of educational research, policy, and practice. We hope that this volume meets this important and challenging undertaking in some small way, moving the needle closer toward a truly complete broad education.

References

Abe, J. A. (2005). The predictive validity of the five-factor model of personality with preschool age children: A nine year follow-up study. *Journal of Research in Personality, 39*, 423–442.

Acrum, R. (1992). *The college application and admission process*. New York: College Board.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9*(1), 9–30.

Bartram, D., Robertson, I. T., & Callinan, M. (2002). Introduction: A framework for examining organizational effectiveness. In I. T. Robertson, M. Callinan, & D. Bartram (Eds.), *Organizational effectiveness: The role of psychology* (pp. 1–10). Chichester, UK: Wiley.

Bogg, T., & Roberts, B. W. (2004). Conscientiousness and health-related behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological Bulletin, 130*(6), 887–919.

Branje, S. J. T., Frijns, T., Finkenauer, C., Engels, R., & Meeus, W. (2007). You are my best friend: Commitment and stability in adolescents' same-sex friendships. *Personal Relationships, 14*(4), 587–603.

Burton, N. W., & Ramist, L. (2001). *Predicting success in college: SAT studies of classes graduating since 1980* (College Board Research Report No. 2001–002). New York: College Board.

Camara, W. J., & Echternacht, G. (2000). *The SAT I and high school grades: Utility in predicting success in college* (RN-10). New York: College Entrance Examination Board, Office of Research and Development.

Camara, W., & Kimmel, E. W. (2005). *Choosing students: Higher education admissions tools for the 21st century*. Mahwah, NJ: Lawrence Erlbaum.

Campbell, J. R., Voelkl, K. E., & Donahue, P. L. (1997). *NAEP 1996 trends in academic progress* (NCES Publication No. 97–985). Washington, DC: U.S. Department of Education.

Casner-Lotto, J., & Barrington, L. (2006). *Are they really ready to work? Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century U.S. workforce*. Washington, DC: Conference Board, Partnership for 21st Century Skills, Corporate Voices for Working Families, Society for Human Resource Management.

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology, 56*, 453–484.

Chamorro-Premuzic, T., & Furnham, A. (2003). Personality traits and academic examination performance. *European Journal of Personality, 17*(3), 237–250.

Crede, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Sciences, 3*, 425–454.

De Fruyt, F., Bartels, M., Van Leeuwen, K. G., De Clercq, B., Decuyper, M., & Mervielde, I. (2006). Five types of personality continuity in childhood and adolescence. *Journal of Personality and Social Psychology, 91*(3), 538–552.

Deary, I. J., Whalley, I. J., & Starr, J. M. (2003). IQ at age 11 and longevity: Results from a follow up of the Scottish Mental Survey 1932. In C. E. Finch, J. M. Robine, & Y. Christen (Eds.), *Brain and longevity: Perspectives in longevity* (pp. 153–164). Berlin: Springer.

Diener, E., & Lucas, R. E. (1999). Personality and subjective well-being. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 213–229). New York: Russell Sage Foundation.

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16*(12), 939–944.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*(1), 4–58.

Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology, 45*(3), 294–309.

Ge, X., & Conger, R. D. (1999). Adjustment problems and emerging personality characteristics from early to late adolescence. *American Journal of Community Psychology, 27*, 429–459.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology, 94*(3), 562–575.

Heckman, J. J., Malofeeva, L., Pinto, R. R., & Savelyev, P. (2007). *The effect of the Perry Preschool program on the cognitive and noncognitive skills of its participants*. Unpublished manuscript, Department of Economics, University of Chicago, IL.

Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *The American Economic Review, 91*(2), 145–149.

Hezlett, S. A., Kuncel, N., Vey, M. A., Ahart, A. M., Ones, D. S., Campbell, J. P., et al. (2001, April). *The effectiveness of the SAT in predicting success early and late in college: A comprehensive meta-analysis*. Paper presented at the annual meeting of the National Council on Measurement in Education, Seattle, WA.

Hough, L. M., & Oswald, F. L. (2008). Personality testing and industrial-organizational psychology: Reflections, progress, and prospects. *Industrial and Organizational Psychology, Perspectives on Science and Practice, 1*(3), 272–290.

James, W. (1890/1981). *The principles of psychology* (F. Burkhardt, Ed., 2 vols.). Cambridge, MA: Harvard University Press.

Jensen-Campbell, L. A., Adams, R., Perry, D. G., Workman, K. A., Furdella, J., & Egan, S. K. (2002). Agreeableness, extraversion, and peer relations in early adolescence: Winning friends and deflecting aggression. *Journal of Research in Personality, 36*, 224–251.

Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology, 87*(3), 530–541.

Kaufman, J. C., & Beghetto, R. A. (2013). In praise of Clark Kent: Creative metacognition and the importance of teaching kids when (not) to be creative. *Roeper Review, 35*, 155–165.

King, M. L. Jr. (1947, January–February). *The purpose of education. The Maroon Tiger*. Retrieved from http://kinfencyclopedia.stanford.edu/encyclopedia/documentsentry/doc_470200

Klimstra, T. A., Hale, W. W., III, Raaijmakers, Q. A. W., Branje, S. J. T., & Meeus, W. H. J. (2009). Maturation of personality in adolescence. *Journal of Personality and Social Psychology, 96*(4), 898–912.

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT for predicting first-year college grade point average* (College Board Research Report Vol. 5, pp. 1–10). New York: College Board.

Kurz, R., & Bartram, D. (2002). Competency and individual performance: Modeling the world of work. In I. T. Robertson, M. Callinan, & D. Bartram (Eds.), *Organizational effectiveness: The role of psychology* (pp. 227–255). Chichester, London: Wiley.

Kyllonen, P. C., Lipnevich, A. A., Burrus, J., & Roberts, R. D. (2014). *Personality, motivation, and college readiness: A prospectus for assessment and development* (ETS Research Report No. ETS RR-14-06). Princeton, NJ: Educational Testing Service.

Kyllonen, P. C., Roberts, R. D., & Stankov, L. (2008). *Extending intelligence: Enhancement and new constructs*. New York: Lawrence Erlbaum Associates.

Lee, J., Redman, M., Goodman, M., & Bauer, M. (2007*). Enhance: Noncognitive assessments for K-12*. Paper presented at the 2007 annual conference of the National Council of Measurement in Education, Chicago, IL.

Lipnevich, A. A., MacCann, C., & Roberts, R. D. (2013). Assessing noncognitive constructs in education: A review of traditional and innovative approaches. In D. H. Saklofske, C. B. Reynolds, & V. L. Schwean (Eds.), *Oxford handbook of child psychological assessment* (pp. 750–772). Cambridge, MA: Oxford University Press.

MacCann, C., Fogarty, G. J., & Roberts, R. D. (2012). Strategies for success in education: Time management is more important for part-time than full-time community college students. *Learning and Individual Differences, 22*(5), 451–458.

Matthews, G., Zeidner, M., & Roberts, R. D. (2006). Personality, affect, and emotional development. In P. A. Alexander & P. H. Winne (Eds.), *Handbook in educational psychology* (2nd ed., pp. 163–186). Mahwah, NJ: Lawrence Erlbaum.

McCrae, R. R., Costa, P. T., Jr., Terracciano, A., Parker, W. D., Mills, C. J., De Fruyt, F., et al. (2002). Personality trait development from age 12 to age 18: Longitudinal, cross-sectional and cross-cultural analyses. *Journal of Personality and Social Psychology, 83*(6), 1456–1468.

National Research Council. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: The National Academies Press.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big Five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology, 93*(1), 116–130.

O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences, 43*, 971–990.

Oswald, F. L., Schmitt, N., Kim, B. H., Ramsay, L. J., & Gillespie, M. A. (2004). Developing a biodata measure and situational judgment inventory as predictors of college student performance. *Journal of Applied Psychology, 89*(2), 187–207.

Ozer, D. J., & Benet-Martinez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology, 57*, 401–421.

Partnership for 21st Century Skills (2006a). Results that matter. 21st century skills and high school reform. Retrieved from. http://www.p21.org/tools-and-resources/publications/204.

Partnership for 21st Century Skills (2006b). State standards for the 21st century. Retrieved from. http://www.p21.org/tools-and-resources/publications.

Partnership for 21st Century Skills (2008). Brief: Policy recommendations on preparing Americans for the global skills race. Retrieved from. http://www.p21.org/toolsand-resources/publications.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322–338.

Prinzie, P., & Dekovic, M. (2008). Continuity and change of childhood personality characteristics through the lens of teachers. *Personality and Individual Differences, 45*(1), 82–88.

Pullmann, H., Raudsepp, L., & Allik, J. (2006). Stability and change in adolescents' personality: A longitudinal study. *European Journal of Personality, 20*(6), 447–459.

Rigol, G. W. (2003). *Admissions decision-making models: How U.S. institutions of higher education select undergraduate students*. New York: The College Board.

Robbins, S., Allen, J., Casillas, A., Peterson, C. H., & Le, H. (2006). Unraveling the differential effects of motivational and skills, social, and self-management measures from traditional predictors of college outcomes. *Journal of Educational Psychology, 98*(3), 598–616.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychology Bulletin, 130*(2), 261–288.

Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology, 58*(1), 103–139.

Roberts, B. W., Kuncel, N., Shiner, R. N., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socio-economic status, and cognitive ability for predicting important life outcomes. *Perspectives in Psychological Science, 2*(4), 313–345.

Roberts, R. D., & Lipnevich, A. A. (2011). From general intelligence to multiple intelligences: Meanings, models, and measures. In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA Educational Psychology Handbook, Volume 2* (pp. 33–57). Washington, DC: APA.

Roberts, R. D., Martin, J., & Olaru, G. (2015). *A Rosetta Stone for noncognitive skills understanding, assessing, and enhancing noncognitive skills in primary and secondary education*. New York: Asia Society and ProExam.

Roberts, B. W., Walton, K., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*(1), 1–25.

Robertson, I. T., Callinan, M., & Bartram, D. (2002). *Organizational effectiveness: The role of psychology*. Chichester: Wiley.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*(2), 262–274.

Shute, V. J., Ventura, M., Kim, Y. J., & Wang, L. (2014). Video games and learning. In W. G. Tierney, Z. Corwin, T. Fullerton, & G. Ragusa (Eds.), *Postsecondary play: The role of games and social media in higher education* (pp. 217–235). Baltimore, MD: John Hopkins University Press.

Stark, S., Chernyshenko, O. S., Drasgow, F., & White, L. A. (2012). Adaptive testing with multidimensional pairwise preference items: Improving the efficiency of personality and other noncognitive assessments. *Organizational Research Methods, 15*(3), 463–487.

Terracciano, A., Costa, P. T., & McRae, R. R. (2006). Personality plasticity after age 30. *Personality & Social Psychology Bulletin, 32*(8), 999–1009.

Tough, P. (2013). *How children succeed: Grit, curiosity, and the hidden power of character*. New York: Houghton Mifflin Harcourt.

Trull, T. J., & Sher, K. J. (1994). Relationship between the five-factor model of personality and Axis I disorders in a nonclinical sample. *Journal of Abnormal Psychology, 103*(2), 350–360.

Wagerman, S. A., & Funder, D. C. (2006). Acquaintance reports of personality and academic achievement: A case for conscientiousness. *Journal of Research in Personality, 41*, 221–229.

Wagner, T. (2010). *The global achievement gap: Why even our best schools don't teach the new survival skills our children need-and what we can do about it*. New York: Basic Books.

Watson, D., Hubbard, B., & Weise, D. (2000). General traits of personality and affectivity as predictors of satisfaction in intimate relationships: Evidence from self- and partner-ratings. *Journal of Personality, 68*(3), 413–449.

Zeidner, M., Roberts, R. D., & Matthews, G. (2002). Can emotional intelligence (EI) be schooled? A critical review. *Educational Psychologist, 37*(4), 215–231.

Ziegler, M., MacCann, C., & Roberts, R. D. (2011). Faking: Knowns, unknowns, and points of contention. In M. Ziegler, C. MacCann, & R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 3–16). New York: Oxford University Press.

Index

A.A. Lipnevich et al. (eds.), *Psychosocial Skills and School Systems in the 21st Century*, The Springer Series on Human Exceptionality,
DOI 10.1007/978-3-319-28606-8