

Unclassified

EDU/CERI/CD(2015)13

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

08-Sep-2015

English - Or. English

**DIRECTORATE FOR EDUCATION AND SKILLS
CENTRE FOR EDUCATIONAL RESEARCH AND INNOVATION (CERI) GOVERNING BOARD**

EDUCATION AND SOCIAL PROGRESS

Framework for the Longitudinal Study of Social and Emotional Skills in Cities

The report was prepared by Oliver P. John, University of California, and Filip De Fruyt, University of Ghent.

For further information, please contact: Koji Miyamoto [Koji.Miyamoto@oecd.org] or Maria Huerta [Maria.Huerta@oecd.org]

JT03381370

Complete document available on OLIS in its original format

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.



EDU/CERI/CD(2015)13
Unclassified

English - Or. English

TABLE OF CONTENTS

1. SCOPE.....3

2. DEFINITION OF SOCIAL AND EMOTIONAL SKILLS4

3. SUMMARY OF THE SOCIAL AND EMOTIONAL SKILLS FRAMEWORK5

 3.1. Working with Others: Two Sets of Interpersonal Skills7

 3.2. Managing Emotions: Regulation Skills8

 3.3. Achieving Goals: Task Performance and Open-Mindedness9

 3.4. Developmental pathways10

4. RATIONALE11

 4.1. Strong Empirical Foundation: The Big Five Personality Domains12

 4.2. High Predictive Power and Comprehensiveness24

 4.3. Malleability40

5. COHERENCE WITH OTHER FRAMEWORKS AND EDUCATIONAL GOALS43

 5.1. Extensions to Other Frameworks45

 5.2. Cultural Sensitivity46

 5.3. Trait Building Blocks and Skills46

6. VALIDATION STRATEGY48

7. POLICY QUESTIONS49

ANNEX A. ALPHABETICAL LIST OF 21ST CENTURY CHARACTERISTICS50

ANNEX B. FEASIBILITY STUDIES: INITIAL METHODOLOGICAL CONSIDERATIONS51

REFERENCES56

Tables

Table 3.1. Proposed Framework: Five Broad Skill Domains and More Specific Socio-Emotional Characteristics within Each Domain Derived from the Literature Review6

Table 4.1. The Big Five OCEAN of Personality Domains: First-Letter Abbreviations, Verbal Labels, Conceptual Definitions, and Three More Specific Facet Traits in Each Domain.....14

Table 4.2. Minimal Set of Three Facets for Each Big Five Domain Based on a Review of Previous Facet Models15

Table 4.3. Socio-Emotional Elaboration of the Big Five: Examples of Self-reported 21st Century Skills.....17

Table 4.4. How Do the Major Domains of the Proposed Framework Relate to Models of Socio-Emotional Learning Domains and Emotional Competencies?33

Table 5.1. Mapping Different Frameworks Proposed for Socio-Emotional Skills and for Goals of Education44

Figures

Figure 3.1. Proposed Social and Emotional Skills Framework5

Figure 3.2. Skills beget skills10

1. SCOPE

The OECD's Longitudinal Study of Skill Development in Cities aims to track the development of children's and adolescents' socio-emotional skills, and to assess how they predict a broad set of outcomes, including educational attainment, employability and labour market position, job performance, health conditions, well-being, interpersonal connectedness, civic engagement and environmental awareness, and crime/safety (OECD, 2015a). These developmental processes will be studied in a variety of cities, each characterized by specific political, socioeconomic and cultural contexts and challenges, forming a natural experiment to examine processes of continuity, change, time course, impacts, and outcomes of socio-emotional skills. The investigated outcomes are directly consequential at the level of the individual, but also impact upon the societal level. Therefore, the findings emerging from such a project should provide much-needed information about the life paths of individuals as they experience particular home, school, and community environments as well as generalizable research findings that have fundamental implications for education policies and practices.

The Longitudinal Study of Social and Emotional Skills in Cities extends prominent OECD projects that include socio-emotional measures, including the Programme of International Student Assessment (PISA). This is done by expanding the coverage of social and emotional assessment domains and introducing an assessment design to identify the growth trajectory of these skills. These ambitious objectives will not only introduce challenges in terms of study design and subsequent research, but also impose extra requirements on the key constructs that will be examined. The proposed Longitudinal Study aims to trace the developmental pathways of the same group of individuals across a long period of time. Instead of assessing the level of educational performance within and across countries and its associations with contextual factors, the Longitudinal Study of Social and Emotional skills in Cities will examine growth trajectories within developing contexts and document how these affect each other in explaining a broad variety of outcomes that are valued by individuals and society in general.

This report presents a conceptual framework of social and emotional skills among school-aged children and adolescents, relying on a review of the most relevant and recent literature. It first defines what socio-emotional skills are and how they have been characterised and structured in different literatures. The next section provides a summary of the proposed framework, beginning with the three core issues identified by the OECD (2015): (a) *Working with Others*, (b) *Managing Emotions*, and (c) *Achieving Goals*. The next section explains the rationale and research background for the particular constructs included in the framework, including predictive power and comprehensiveness, malleability, and temporal stability. The next section examines the coherence of the proposed framework with other frameworks and educational goals. The report ends with two briefer sections, one outlining strategies for validating the proposed framework over the next several years and the other outlining policy questions.

The proposed framework has a strong empirical foundation. It is based on a large number of psychometric analyses of micro-data conducted by psychologists across countries, languages and cultures. This framework may not be consistent with similar social and emotional skills frameworks proposed by educators due to the differences in the methodologies employed. For instance, some of these frameworks have been derived by synthesizing other existing frameworks, complemented by qualitative interviews with diverse stakeholders including teachers, parents and employers.

The proposed framework has a strong empirical foundation. It is based on a large number of psychometric analyses of micro-data conducted by psychologists across countries, languages, and cultures. This framework may not be consistent with similar social and emotional skills frameworks proposed by educators due to the differences in the methodologies employed. For instance, some of these frameworks have been derived by synthesizing other existing frameworks, complemented by qualitative interviews with diverse stakeholders including teachers, parents and employers.

2. DEFINITION OF SOCIAL AND EMOTIONAL SKILLS

Interest in social and emotional skills (e.g., goal-setting, perseverance, optimism, emotional control, gratitude, social intelligence, curiosity, etc.) has a long history. Education researchers, school administrators, teachers, parents, and even the children themselves have long been aware that education involves interactions among people; people are not only inherently social creatures but also experience and express a wide range of emotions. In other words, most schools are places that are both intensely social and intensely emotional. Previous OECD publications have shown that the way in which students, parents and teachers navigate these social and emotional processes can have powerful consequences for a multitude of important life outcomes (e.g., Kautz, et al., 2014; OECD, 2015).

In recent years, socio-emotional characteristics have also been referred to as a key component of 21st century or *employability skills* (e.g., Trilling and Fadel, 2009) because they include a set of competencies considered increasingly crucial for individuals' development, employment, and healthy functioning in current and future societies (National Academy of Sciences, 2012). As individuals and jobs become increasingly interconnected, complex, and collaborative, socio-emotional characteristics are expected to become ever more important. Many different 21st century skills have been proposed over the years (see Annex 1 for a list of more than 160 individual skills described in Trilling and Fadel, 2009 and Fadel, 2014). They include such concepts as abnegation and altruism, engagement and enthusiasm, innovation and inquisitiveness, self-discipline and self-control, stability and tranquillity, and many more.

Many of these characteristics have also interested psychologists, who have studied them under the broad rubric of "personality traits" (John, 1990). As described below, extensive research has shown that these personality concepts can be organized into five broad and relatively independent and distinct domains of individual differences in thinking, feeling, and behaving, often referred to as the *Big Five* or the *Five Factor Model*. Moreover, as will be reviewed below, modern research has shown that traits are not in-born and fixed; in contrast to popular views, personality traits develop through the interplay of personal and environmental factors (i.e., learning) and they show considerable plasticity, especially during childhood and adolescence.

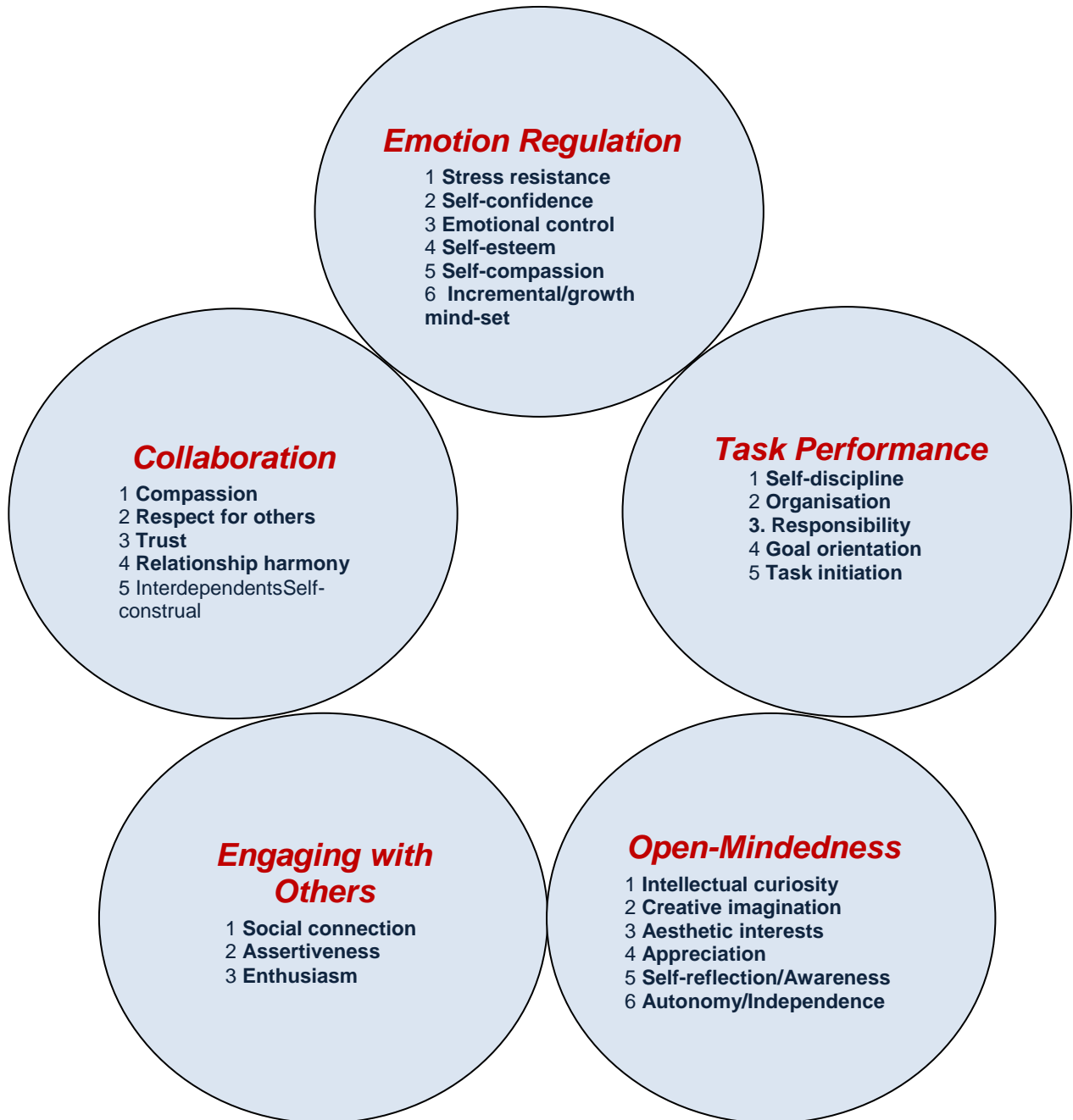
At the same time, developmental psychologists and educational researchers have studied how socio-emotional learning can be improved through school-based interventions. They have focused on specific interventions, and their particular contexts and unique outcomes. These researchers have been less interested in developing a single, generally accepted model that could organize the hundreds of individual social and emotional skills into one coherent taxonomy, like the table of elements in chemistry or the taxonomy of the animal kingdom in biology. Instead, they have developed multiple heuristic models. For example, Elias et al. (1997) proposed 6 (or 7) major domains of socio-emotional learning; Durlak et al. (2011) proposed 5 domains; and Saarni (2011) proposed 8. These three models differ in the number of major skill constructs they include, and they are not fully consistent with each other. Neither of them is comprehensive; they emphasize different aspects of the construct space. Section 5 of this report describes in more detail how these frameworks relate to, and differ, from each other.

The OECD (2015) defines social and emotional skills as: "*individual capacities that (a) are manifested in consistent patterns of thoughts, feelings and behaviours, (b) can be developed through formal and informal learning experiences, and (c) influence important socioeconomic outcomes throughout individual's life*". This definition captures the essential features of social and emotional skills that are reflected in all the constructs proposed in the next sections as well as in the literature of 21st Century skills, personality psychology, developmental psychology, and social and emotional learning.

3. SUMMARY OF THE SOCIAL AND EMOTIONAL SKILLS FRAMEWORK

Figure 1 summarizes the proposed Social and Emotional Skills framework that builds on the vast conceptual and empirical literature that pertains to social and emotional skills, which will be described in the subsequent sections of this report. This section briefly describes this framework and the next section justifies the choice of the five broad domains and the corresponding lower-level facets.

Figure 3.1. Proposed Social and Emotional Skills Framework



The OECD’s Longitudinal Study of Skill Development in Cities specified three domains of individual functioning that are of particular interest: Managing Emotions, Working with Others, and Achieving Goals. The conceptual framework should be able to account for the most important socio-emotional skills that are relevant to these three domains. However, these three domains are not themselves elements of the empirical framework because they are formative latent constructs: they are called formative because they are (like the Consumer Price Index) formed for practical or descriptive purposes and often identified through consensus opinion or traditions in a field. Formative latent variables have been described as a “stew”—a mixture of more basic elements that might or might not be related. For example, “Working with Others” mixes together more basic elements such as helping others, leading others, and following others that are psychologically distinct from each other and empirically rather different.

These more basic elements are called reflective latent variables because they reflect the essence or commonality of the various specific skills that can be measured. In psychometric models, these reflective latent variables are also called factors because they are often discovered through the statistical method of factor analysis. As will become apparent in Section 5 of this report, the elements of the framework proposed here can be linked systematically to formative frameworks that articulate and advocate particular goals for education. Indeed, it has been argued that “The various lists of 21st century skills that have been proposed to date are formative variables, identified by consensus opinion” (National Academy of Sciences, 2012, p. 27).

What are these more basic, reflective latent constructs? At the highest level of abstraction, socio-emotional skills constructs can be divided into five broad domains as shown in Figure 3.1. However, these five distinctions are at a very abstract level, and each of the five domains has therefore been subdivided into several more narrowly defined, specific constructs labeled facets. Table 3.1 lists a brief label for each of these facets, provides a short definition, and (if available) illustrates the concepts with an example item in parentheses.

Table 3.1. Proposed Framework: Five Broad Skill Domains and More Specific Socio-Emotional Characteristics within Each Domain Derived from the Literature Review

<p>1. Engaging with Others</p>
<p>1 Social approach and connection: Able to approach others, both friends and strangers, initiating and maintaining social connections; skilled at teamwork, including communication and public speaking skills (Is outgoing, comfortable around people)</p> <p>2 Assertiveness (or courage): Able to voice opinions, needs, and feelings, and exert social influence; capacity to assert own will to accomplish goals in the face of opposition, such as speaking out, taking a stand, and confronting others if needed; courage (Takes on leadership roles)</p> <p>3 Enthusiasm: Passion and zest for life; approaching daily life with energy, excitement, and spontaneity (Is full of energy, shows enthusiasm)</p>
<p>2. Collaboration: “Tending and Befriending” Others</p>
<p>1 Compassion: Kindness and caring for others stems from perspective taking and empathic concern for their well-being, and leads to valuing and investing in close relationships (Considerate and kind to everyone)</p> <p>2 Respect for others (politeness): Treating people with respect and politeness, the way oneself would like to be treated, according to notions of fairness, justice, and tolerance (Is respectful; treats others with respect vs. breaking rules)</p> <p>3 Trust: assuming that others generally have good intentions and forgiving those that have done wrong; avoid being harsh and judgmental, giving people another chance (Assumes the best about people)</p> <p>4 Relationship harmony: Living in harmony with others and valuing interconnectedness among all people; being inclusive of others who have different backgrounds, customs, and beliefs (It is important to me to respect decisions made by the group)</p> <p>5 Interdependent self-construal: Experiencing self as part of a collective, interconnected and inseparable from important groups, such as family (I feel my fate is intertwined with the fate of those around me)</p>

<p>3. Emotion Regulation</p> <p>1 Stress resistance: Effectiveness in modulating anxiety and response to stress; untroubled by excessive worry and able to calmly solve problems (Is relaxed, handles stress well)</p> <p>2 Self-confidence: Positive and optimistic expectations for self and life; anticipates success in actions undertaken; a “can-do” mind-set (Feels secure, comfortable with self)</p> <p>3 Emotional control: Effective strategies for regulating temper, anger, and irritation; able to maintain tranquillity and equanimity in the face of frustrations; not moody or volatile (Keeps their emotions and temper under control)</p> <p>4 Self-esteem: Acceptance and positive evaluation of oneself (I am a person of worth.)</p> <p>5 Self-compassion: Taking a mindful, kind, and accepting approach towards oneself, rather than being overly critical or self-blaming (When going through a hard time, I give myself the caring and tenderness I need)</p> <p>6 Incremental (or Growth) mind-set: Believing that things are changeable; that humans can improve, learn, and grow; and that effort will improve one’s personal future (When bad things happen, I think about ways to make things better, rather than “what’s wrong with me”)</p> <p>7 Fear of happiness: Beliefs and worry that happiness will lead to bad outcomes (I prefer not to be too joyful, because usually joy is followed by sadness)</p>
<p>4. Task Performance</p> <p>1 Self-discipline: Grit, perseverance, and effortful control are related concepts that involve concentration skills: the ability to focus attention on the current task and avoid distractions in order to achieve personal goals (Is efficient, gets things done)</p> <p>2 Organisation: Organisational skills are critical for planning and executing plans to reach longer-term goals (Keeps things neat and tidy)</p> <p>3 Responsibility: Time management, punctuality, and honouring commitments are critical to reliability and consistency, and engender trustworthiness (Is reliable, can always be counted on)</p> <p>4 Goal orientation: Setting high standards for oneself and working hard to meet them, as illustrated by a strong “work ethic”, consistent effort, and high levels of productivity (Wants to be excel at everything s/he does)</p> <p>5 Task initiation: Ability to get started on a task or goal, rather than engaging in prolonged procrastination (Leaves difficult tasks for later vs. tackles them immediately)</p>
<p>5. Open-Mindedness: Interest and devotion to matters of the mind</p> <p>1 Intellectual curiosity: Interest in ideas and love of learning, understanding, and intellectual exploration; an inquisitive mind-set (Likes to think, play with ideas)</p> <p>2 Creative imagination : Generating novel ways to do or think about things through tinkering, learning from failure, insight, and vision (Is original, comes up with new ideas)</p> <p>3 Aesthetic interests: Valuing art and beauty that may be experienced or expressed through music, writing, visual and performing arts, and other forms of self-actualization (Is fascinated by music, art, or literature)</p> <p>4 Appreciation: Valuing and noticing the environment, living in harmony with nature, spirituality, awe, and reverence</p> <p>5 Self-reflection/Awareness of inner experiences: Awareness of inner processes and subjective experiences, such as thoughts and feelings, and the ability to reflect about and articulate such experiences (meta-cognition)</p> <p>6 Autonomy/Independence of judgment and self-construal: Thinking for yourself; grounding beliefs, attitudes, and values on a critical analysis through independent thought (I enjoy being unique and different from others in many respects)</p>

3.1. Working with Others: Two Sets of Interpersonal Skills

We begin with the constructs most relevant to the OECD domain “Working with Others.” Indeed, schools are intensely social settings, with students, teachers, parents, and school administrators all interacting with each other and forming relationships. Thus, schools provide rich environments for acquiring and practicing many kinds of social skills. These fall into two distinct sets within the framework. One set, labelled here *Engagement with Others*, captures the basic interpersonal direction of the individual, *towards* engagement and interaction with others, as contrasted with avoidance or withdrawal *away from* interpersonal contact. The other set, *Collaboration*, captures the quality of the interactions and relationships: is the child able to construe others as likely friends and sources of pleasure and support that can be trusted and loved (amity), or as adversaries (enmity)?

Within the superordinate domain *Engagement with Others*, the research literature agrees on three further distinctions. The first facet is *Social approach and connection*: learning to approach others and initiate and maintain connections with them is a critical learning task during childhood and adolescence. This involves practicing communication and public speaking skills (e.g., being able to tell a joke or story to a group of friends or give a short presentation in front of the class), which are critical to all kinds of team work. The second facet is Assertiveness (or courage), which involves children finding their “voice” and practicing it by speaking out, taking a stand, and confronting others if needed to achieve goals or meet their needs. The third facet is Enthusiasm, defined as passion and zest for life, approaching every day with energy, excitement, and spontaneity. These three facets are clearly interpersonal and thus most relevant to the OECD theme of “Working with Others”; however, these social and communication skills may also help children with the third OECD theme, namely Achieving Goals.

The second interpersonal domain in the framework, *Collaboration*, involves perceiving and treating others as friends, as expressed by the idea that others exist for our “Tending and Befriending”. Again, three core facets have been identified. *Compassion* refers to kindness and caring for others that stems from perspective taking and empathic concern for their well-being; *Respect for others* involves treating people with respect and politeness, the way oneself would like to be treated; and *Trust* involves the belief that others generally have good intentions and forgiving those that have done wrong. In addition to these three core facets, cross-cultural research reviewed below suggested that Western notions of Collaboration constructs may need to be supplemented with other ways of experiencing the connection between self and other: one additional construct is *Living in harmony* with others and valuing interconnectedness among people; the other is *Interdependent Self-Construct* defined as experiencing the self as part of a collective, interconnected and inseparable from important groups, such as one’s family.

3.2. Managing Emotions: Regulation Skills

The first set of constructs reviewed so far involved characteristics that are primarily social or *inter*-personal. However, learning environments in general, and schools in particular, are also rich with emotions, which are primarily *intra*-personal phenomena. Researchers generally agree that the way students learn to manage their emotions is of critical importance for their concurrent and subsequent adjustment and well-being. Psychologists have long recognized three basic and universal negative emotions that can potentially undermine students’ well-being: fear, sadness, and anger (Ekman, 1972). Fear (and anxiety) arises when students are stressed by danger or uncertainty; anxiety can lead to a cascade of negative consequences through avoidance behaviour when students try to control their anxiety by avoiding the situations that make them anxious. Sadness (and depression) occurs when students experience disappointments, failures, and losses; sadness is sapping the individual of energy and can lead to social withdrawal (i.e., lower engagement). Anger typically arises from frustrations and the perception that one is not getting what one wants or deserves (i.e., when our actions or wishes are blocked by specific others, by rules, etc.). Thus, effectively regulating these negative emotions is of considerable importance, and emotion researchers (e.g., Saarni, 1999) have repeatedly called for teaching students emotion regulation skills in school. In the proposed framework, the broad domain of Emotion Regulation specifies three core facets. *Stress Resistance* reflects effectiveness in modulating anxiety and stress responses. *Self-confidence* includes positive expectations for self and life, anticipating success, and a “can-do” mind-set that helps buffer the child from sadness and depression. And *Emotional Control* refers to effective strategies for regulating temper, anger, and irritation that help maintain tranquillity and equanimity in the face of frustration. In addition to these three core facets, the framework includes a construct that has been widely studied as an emotion-protective factor, namely *Self-esteem*, defined as acceptance and positive evaluation of oneself. Three additional constructs were added on the basis of research on cognitive factors and cross-cultural variation. Beliefs about controllability have been widely studied, and the most promising recent intervention research points to the power of mind-sets; the *Incremental (or Growth) Mind-Set* involves

beliefs that things are changeable, that humans can improve, learn, and grow, and that effort will improve one's personal future. The other construct was suggested by recent research on mindfulness and emotion regulation: *Self-compassion* involves taking a mindful, kind, and accepting approach towards oneself when making mistakes or experiencing failure or set-backs, rather than being overly critical or self-blaming; this approach to self-care tends to improve emotional functioning. Finally, cross-cultural research has demonstrated substantial differences among countries and cultural groups in the ways emotions (especially happiness, pride, and love) are experienced and regulated; thus the framework includes a concept that has shown substantial cultural differences, namely *Fear of Happiness*, which involves beliefs and worries that happiness will lead to bad outcomes (e.g., "I prefer not to be too joyful, because usually joy is followed by sadness").

3.3. Achieving Goals: Task Performance and Open-Mindedness

Finally, even though some conceptions of socio-emotional functioning do not include skills related to goal achievement, the OECD did emphasize that achieving goals ought to be addressed in the longitudinal study. Indeed, considerable research has accumulated evidence that there are two distinct sets of skills that predict achievement behaviours (e.g., such as completing homework, studying regularly, and school attendance) as well as achievement outcomes (e.g., scores on standardized achievement tests and grades).

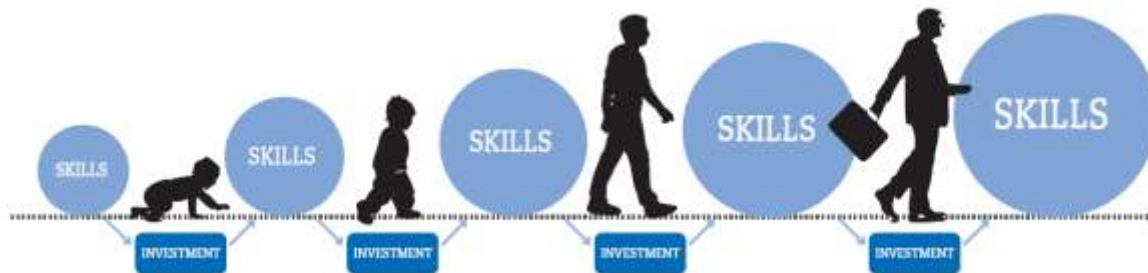
The first set involves *Task Performance* (labelled Conscientiousness in the Big Five personality work) and includes three commonly studied core facets. First, a number of constructs defined and studied separately in the research literature (e.g., grit, perseverance, effortful control), are in fact related concepts that all involve what here is called *Self-discipline*, namely concentration skills and the ability to focus attention on the current task and avoid distractions in order to achieve personal goals. The second facet is *Organization*: organizational skills are critical for planning and executing plans to reach longer-term goals. The third facet, *Responsibility*, also has some interpersonal implications because time management skills, punctuality, and honouring commitments are critical to being perceived as reliable and consistent and engender trustworthiness. In addition, we included two other facets to ensure adequate coverage of this important domain. *Goal orientation* involves setting high standards for oneself and working hard to meet those standards (e.g., work ethic), and *Task initiation* involves the ability to get started on a task or goal immediately, rather than engaging in prolonged procrastination.

The second set is here called *Open-Mindedness: Interest and devotion to matters of the mind*. The three most commonly studied facets are *Intellectual Curiosity* (defined as a passionate interest in ideas and the desire to learn and understand, and intellectual exploration); *Creative Imagination* (Generating novel ways to do or think about things through tinkering, learning from failure, insight, and vision); and Aesthetic Interests (Valuing art and beauty that may be experienced or expressed through music, writing, visual and performing arts, and other forms of self-expression and self-actualization). In addition, research on positive psychology has emphasized the importance of a spiritual connection to nature and its meaning-making potential, which led us to add the concept of *Appreciation* (Valuing and noticing the environment, living in harmony with nature, spirituality, awe, and reverence). Research on cognitive approaches and meta-cognition suggested that awareness of mental processes may be important, while research on emotional competence suggested that awareness of emotional experience may be important; thus, an additional facet was included to capture skills related to *Self-reflection, Introspection, and Awareness of inner experiences* (Awareness of inner processes and subjective experiences, such as thoughts and feelings, and the ability to reflect about and articulate such experiences; meta-cognition). The final facet shown in Table 3.1 was identified in research on cultural differences between Western and East Asian countries and involves the relative importance of *Autonomy and Independence of Judgment and Self-construal* (Thinking for yourself; grounding beliefs, attitudes, and values on a critical analysis through independent thought) versus following societal traditions and conventions.

3.4. Developmental pathways

Social and emotional skills are assumed to develop progressively over time, building on skills already accumulated as well as learning inputs (or, investments) from parents, teachers and the community (Figure 3.2).

Figure 3.2. Skills beget skills



Source: OECD (2015)

Past research conducted by developmental psychologists and economists are broadly consistent with this model (see OECD, 2015 for some evidence on this). However, little is known about the precise nature of the dynamic formation of social and emotional skills.

Policy-makers, parents and teachers would benefit from knowing the optimal learning inputs (e.g. parenting, specific curricular activities) that are conducive to social and emotional development during each period of the child’s development. They would also be interested in learning about the degree of malleability of social and emotional skills and the optimal mix of social and emotional skills during each developmental period. This information would help education stakeholders prepare a sequence of learning environments that would allow children to accumulate sufficient levels of social and emotional skills before they enter adulthood. There is a need to develop better data and analyses to disentangle such a complex nature of dynamic skill formation. The proposed Longitudinal Study of Social and Emotional skills in Cities is expected to contribute to this much-needed evidence.

4. RATIONALE

The proposed social-emotional skills framework was developed according to the following set of **general principles**. More detailed information will be provided later in this section.

Strong empirical foundation

The proposed framework should be based on a strong empirical foundation. Recent advances in the field of personality psychology have identified the *Big Five personality taxonomy* (John, 1990), also referred to as the Five-Factor Model (FFM; McCrae & Costa, 1996) as the most empirically compelling model that can serve as the starting point of developing a more comprehensive framework that meets the needs of the Longitudinal Study of Social and Emotional skills in Cities. The core dimensions of this model are usually referred to as Extraversion (vs. Introversion), Agreeableness (vs. Antagonism), Conscientiousness (vs. Lack of Direction), Emotional Stability (vs. Neuroticism), and Openness to experience (vs. Closed-mindedness). As literally thousands of concepts have been proposed to describe and explain individual differences in personality functioning, the emergence of this integrative taxonomy and its general acceptance in the 1990s has brought order and coherence into the field and has led to a remarkable surge in research productivity and findings (John, Naumann, & Soto, 2008). Although the Big Five taxonomy was initially derived from research on adults, it has been well-documented that the Big Five are suitable to describe personality differences from childhood to old age (e.g., De Fruyt & De Clercq, 2014; Soto, John, Gosling, & Potter, 2008, 2011). The proposed framework should also be **cross-culturally relevant** given the diverse countries and cultures that the Longitudinal Study of Social and Emotional skills in Cities is scheduled to cover.

High predictive power and comprehensiveness

Given the ambition of this study is to explain and predict a variety of life outcomes, a comprehensive model of skills will have to be assessed across the developmental trajectory. The Big Five dimensions broadly capture the underlying core qualities of the individual—typical patterns of thoughts, feelings, and behaviours—that drive their lifetime success in life, and thus provide a parsimonious and highly efficient summary. However, such a parsimonious model with few concepts is by necessity very broad and limited in predicting specific outcomes (e.g., Hampson, John, & Goldberg, 1986). Thus, the broad level of the Big Five domains is unlikely to represent the most appropriate level of assessment for all the goals of this project, including understanding growth trajectories, examining the impact of different sorts of environmental factors, or explaining consequential outcomes. The measurement model proposed here will hence need further specification at a more fine-grained level, ultimately with about 3 to 5 facets comprising each of the broad Big Five domains. The final set of non-cognitive constructs should thus enable researchers to follow individuals across both broad and more specific levels of the personality trait hierarchy. Moreover, traits measured at the more specific facet level have the potential to be combined into skill compounds, better reflecting the complex nature of socio-emotional skills. A construct like Grit (Duckworth, Peterson, Matthews, & Kelly, 2007), for example, may be conceptualized as the aggregate of particular facets of Conscientiousness and possibly Extraversion (i.e., the enthusiasm facet here). As explained below, although there is a widespread consensus on the Big Five, there has been less research and thus agreement on the structure of the lower-order level of facet traits (e.g., John et al., 2008, Table 4.3). A crucial task for future work will hence be to identify those facets that are most likely to predict key outcomes of interest.

Malleability

The socio-emotional skill battery will have to include measures that are sensitive to detecting reliable changes, including normative change pattern (Roberts, Walton, & Viechtbauer, 2006; Roberts, Wood, & Smith, 2005) that so far have been studied much more widely in adults than in younger age groups. For example, researchers have hypothesized increases in conscientiousness, agreeableness and emotional stability to begin in mid-to-late adolescence, but evidence has been mainly come from large studies that used cross-sectional (not longitudinal) designs studies (e.g., Soto, John, Gosling, & Potter, 2011). Moreover, theory also suggests that individual change patterns will be important, as the onset and timing of age-related changes will differ across individuals; again, supporting evidence is available from adult samples but there are only a few longitudinal studies (De Fruyt et al., 2006; Prinzie et al., 2005) that have followed the same children or adolescents over longer periods of time, and all have been done in Western societies. Thus, the present design will include a variety of cultures and multiple assessment points to be able to chart non-linear forms of growth, and to distinguish latent groups of persons following specific developmental trajectories, beyond or deviant from normative change (De Fruyt & Van Leeuwen, 2014).

Temporal stability

The socio-emotional skills presented in the framework tend to be “manifested in consistent patterns of thoughts, feelings and behaviours”; that is, they show sufficient temporal consistency over short time intervals (e.g., one month) to allow for reliable measurement. Indeed, if a characteristic was largely determined by temporary circumstances and showed no temporal stability, then that characteristic would not be considered a skill but a transient state or behaviour. Empirical evidence suggests that many of the lower-order facets presented in Table 3.1 are likely to satisfy the condition of temporal stability to permit reliable measurement (see Almlund, et al., 2011; and Kautz et al., 2014).

4.1. Strong Empirical Foundation: The Big Five Personality Domains***The Empirical Development of a Taxonomy for Personality Attributes***

Until the late 1980s, the Big Five personality dimensions, now seemingly ubiquitous, were hardly known (see John et al., 2008). Researchers and practitioners were faced with a bewildering array of personality scales from which to choose, with little guidance and no organizing theory or framework at hand. What made matters worse was that scales with the same names might measure concepts that were quite different, and scales with different names might measure concepts that were quite similar. Systematic accumulation of research results and communication across researchers was impossible amidst this cacophony of competing concepts and scales.

In the 1970s and 1980s, researchers at the University of California, Berkeley, for example, measured personality with only two concepts (e.g., Block’s two dimensions of Ego-resilience and Ego-control), with the four scales on the Myers-Briggs Type Indicator (MBTI) as well as the 20 scales on the California Psychological Inventory. Many personality researchers were hoping to be the one who would discover the “true” structure that all others would then adopt, thus transforming the fragmented field into a community speaking a common language. However, we now know that such an integration was not to be achieved by any one researcher or any one theoretical perspective. As Allport once put it, “Each assessor has his own pet units and uses a pet battery of diagnostic devices” (1958, p. 258).

The field of personality psychology lacked a descriptive model (or taxonomy), of its subject matter. One of the central goals of scientific taxonomies is the definition of overarching domains within which large numbers of specific instances can be understood in a simplified way (John, 1990). Thus, in personality, a generally accepted taxonomy would specify domains of related personality characteristics

that researchers could study, rather than studying separately each of the thousands of particular attributes that make human beings individual and unique. Moreover, such a taxonomy can help accumulate, summarize, and communicate empirical findings by offering a standard nomenclature.

After decades of research, personality psychology has finally agreed upon a general taxonomy of personality traits, the so-called Big Five personality domains. Table 4.1 presents a brief definition and summary of each domain. Although there remain some critics and contrarians (e.g., Block, 2010), there is now considerable agreement about the Big Five in the personality literature (for a review, see John et al., 2008).

However, there is less agreement about the more specific components or *facets* that define each Big Five domain at a lower level of abstraction. As summarized in Table 4.2, Soto and John (2015) recently reviewed the major facet models, which range from a minimum of only 2 facets per Big Five domain (DeYoung et al., 2006) to 6 facets per domain (Costa and McCrae, 1992). Considering the communalities among the existing models, three major facets emerged as core themes that virtually all of the different models had in common. Soto and John thus concluded that these three facets are widely accepted as important in the field and thus the most worthy to focus on. For example, each of the three facets for Negative Affect reflects one of the three basic emotions of anxiety/fear, sadness, and anger that are commonly accepted as fundamental and universal by emotion researchers (Ekman, 1972).

These three lower-order facet traits are also listed in Table 4.1 with the definition for each Big Five domain. In other words, the Big Five model specifies more than 5 basic constructs because it is hierarchical; instead, each of these five big (i.e., broad) domains consists of several more narrowly defined constructs at a lower level of abstraction (see John et al., 2008, Table 4.3).

Finally, the Big Five dimensions do not represent a particular theoretical perspective but were derived from analyses of the natural language terms people use to describe themselves and others, much like the 21st Century Skills concepts listed in Annex 1. Rather than replacing previous systems, the Big Five taxonomy can serve an integrative function: it represents the various and diverse earlier systems of personality description within a single, common framework (see John et al., 2008, Table 4.1). Could the same taxonomy be applied to bring some order to the myriad number of diverse socio-emotional skills constructs?

Table 4.1. The Big Five OCEAN of Personality Domains: First-Letter Abbreviations, Verbal Labels, Conceptual Definitions, and Three More Specific Facet Traits in Each Domain

<p>O: Openness, Originality, Open-mindedness</p> <p>The attributes in this domain describe the breadth, depth, originality, and complexity of an individual’s mental and experiential life. Facet traits include <i>Intellectual curiosity, Imagination/creativity</i>, as well as <i>Aesthetic and spiritual awareness</i>.</p>
<p>C: Conscientiousness, Constraint, Control of Attention</p> <p>The attributes in this domain describe socially prescribed, effortful self-control that facilitates task- and goal-directed behaviour, such as thinking before acting, delaying gratification, following rules and norms, and planning, organizing, and prioritizing complex and long-term tasks. Facet traits include <i>Self-discipline, Orderliness, and Reliability</i>.</p>
<p>E: Extraversion, Energy, Enthusiasm</p> <p>The characteristics in this domain involve an energetic approach toward the social and material world and include more specific facet traits such as <i>sociability, assertiveness, and positive activity</i>.</p>
<p>A: Agreeableness, Altruism, Affection</p> <p>These characteristics contrast a pro-social and communal orientation towards others with antagonistic or antisocial tendencies, and include facet traits like <i>compassion, respect/politeness, and trust</i>.</p>
<p>N: Negative Emotionality, Nervousness, Neuroticism</p> <p>These characteristics contrast emotional stability, confidence, and even-temperedness with the tendency to experience negative emotions, such as feeling <i>Anxious/nervous, Sad/depressed, or Angry/frustrated</i>.</p>

An Exploratory Pilot Study of Self-reported Socio-emotional Skills: Elaborating the Socio-Emotional Content of the Five Personality Factors

Consider the 21st century attributes listed in Annex 1. To identify links between these socio-emotional skills and the Big Five model in adulthood, John and Mauskopf (2015) conducted a pilot study of self-rated socio-emotional skills and personality characteristics. Specifically, 452 volunteers were presented with an on-line questionnaire that included both socio-emotional skill items and the items from the standard Big Five Inventory. Correlational and factor analyses of these self-ratings showed that the socio-emotional skill list contained much content that was related to the Big Five dimensions. Table 4.3 shows examples of the 21st Century Skills items that had the strongest correlations with the Big Five personality dimensions.

Table 4.2. Minimal Set of Three Facets for Each Big Five Domain Based on a Review of Previous Facet Models

Facets necessary to each Big Five	minimally represent	NEO PI-R (Costa & McCrae, 2008)	AB5C (Goldberg, 1999; Hofstee et al., 1992)	Lexical subcomponents (Saucier & Ostendorf, 1999)	Big Five aspects (DeYoung et al., 2006)
<i>Extraversion</i>					
Social connection		Gregariousness	Gregariousness	Sociability	Enthusiasm
Assertiveness		Assertiveness	Assertiveness	Assertiveness	Assertiveness
Enthusiasm		Activity	—	Activity-Adventurousness	Enthusiasm
<i>Agreeableness</i>					
Compassion		Altruism	Understanding	Warmth-Affection	Compassion
Respect for others		Compliance	Cooperation	Gentleness	Politeness
Trust		Trust	Pleasantness	—	—
<i>Conscientiousness</i>					
Organization		Order	Orderliness	Orderliness	Orderliness
Self-discipline		Self-Discipline	Efficiency	Industriousness	Industriousness
Responsibility		Dutifulness	Dutifulness	Reliability	—
<i>Negative Emotionality</i>					
Anxiety		Anxiety	Toughness (R)	Emotionality	Withdrawal
Depression		Depression	Happiness (R)	Insecurity	Withdrawal
Volatility		Angry Hostility	Stability (R)	Irritability	Volatility
<i>Openness to Experience</i>					
Aesthetic interests		Aesthetics	Reflection	—	Openness
Intellectual curiosity		Ideas	Intellect	Intellect	Intellect
Creative imagination		—	Ingenuity	Imagination-Creativity	—

Note. Based on Soto and John (2015), who used the Big Five labels Negative Emotionality (rather than the older Neuroticism) and Open-mindedness (rather than Openness)

However, despite many clear similarities, there were some noteworthy differences when the socio-emotional item content was considered. First, corresponding to Trilling and Fadel's (2009) emphasis on positive characteristics and interpersonal strengths needed for interlinked *Collaboration*, the largest factor was interpersonal and related to the Big Five domain described earlier as "A" (i.e., Altruism, Affection, Agreeableness). This factor attracted by far the largest number of 21 century skills items. The emphasis here was more explicitly on genuine mutuality and reciprocal exchange. In addition to standard facets of compassion, respect/politeness, and trust, several additional items suggested a potentially new facet, captured well by the item *Living in harmony with others*, along with other interpersonal skills related to interconnectedness and inclusiveness.

As shown in Table 4.3, the second 21st Century Skill factor may be described as *Task Performance*. It was defined by a large number of attributes and was conceptually quite similar to the personality domain of Conscientiousness. In addition to items representing Self-discipline, Organisation, and Responsibility facets, another group of socio-emotional skill items suggested a facet best described as Goal orientation, which highlights positive motivational characteristics like effort, work ethic, and productivity. Again, the strength-based, positive-psychology origin of the 21st Century Skills items rounds out this version of the Big Five "C" factor in a more substantial way.

The third factor, *Emotion Regulation skills*, also highlights positive strengths, whereas the traditional personality literature had focused on the negative, distressing emotions defining the other pole of this dimension. In terms of the more specific facets, we see the opposite of *anxiety*, worry, and avoidance, namely: self-esteem and self-confident, decisive tackling of tough problems. Instead of *sadness* and depression, the focus is on happiness and cheerful optimism. Instead of *anger*, temper, and frustration, there is equanimity, tranquillity, and balance. An additional aspect of emotional strength here is self-compassion (and self-kindness), similar to Kristin Neff's (2007) constructs anchored in research on mindfulness that help the individual avoid self-blame and respond to failures and set-backs in a measured, self-accepting, and normalizing way.

The fourth socio-emotional factor emphasized skills that allow the individual to constructively and joyfully engage with others in their social world. Even though "only" 4th largest in number of items, this factor reminds us of the great importance of positive engagement with the school environment for children and adolescents. Interestingly, the cluster of items related to the Assertiveness facet is enriched by items highlighting proactive strengths, such as leadership and charisma, courage, and the willingness to take a stand. The Enthusiasm facet is enriched by positive-psychology concept like passion and zest for life, spunk and spontaneity, as well as playfulness and humour. We hope there is a place for these positive characteristics in our schools, and that our schools will encourage and nurture these positive emotions in our children.

Table 4.3. Socio-Emotional Elaboration of the Big Five: Examples of Self-reported 21st Century Skills

Factor I: Collaboration (Related to Big Five Agreeableness)

- 1 Compassion, care, cooperation, kindness
- 2 Respect for others, empathy, tolerance, fairness
- 3 Trust, forgiveness, gratitude, appreciation of others
- 4 Living in harmony with others, interconnectedness, inclusiveness

Factor II: Task Performance (Related to Big Five Conscientiousness)

- 1 Self-discipline, focus, perseverance, self-control at school, grit
- 2 Organization, diligence, precision
- 3 Dependability, reliability, consistency, trustworthiness
- 4 Goal orientation, motivation, work ethic, effort, productivity

Factor III: Emotion Regulation (Related to low levels of Negative Emotionality)

- 1 Self-confidence, self-esteem, decisiveness, tackling tough problems
- 2 Cheerfulness, happiness, optimism
- 3 Tranquillity, balance, stability, equanimity (composure and even temper in difficult situations)
- 4 Self-compassion, self-kindness (being positive and understanding towards yourself when you suffer, fail, or feel inadequate)

Factor IV: Engagement with Others (Related to Extraversion)

- 1 Social connection, teamwork, social awareness, public speaking
- 2 Assertiveness, leadership, courage, charisma, speaking out/taking a stand, bravery
- 3 Enthusiasm, passion, zest, inspiration, spunk, spontaneity, playfulness, humour

Factor V: Open-mindedness: The Inquiring Mind (Related to Openness)

- 1 Curiosity, inquisitiveness, willingness to try new ideas, receptivity
- 2 Innovation, vision, insight, tinkering (inventing), learning from mistakes and failures, excitement of creating something new
- 3 Appreciating beauty in the world, living in harmony with nature, spirituality, mindfulness, existentiality, awe, wonder, reverence
- 4 Self-reflection, self-awareness, consciousness, self-actualization, authenticity

Note. Based on John and Mauskopf (2015).

The fifth socio-emotional domain was, *Open-Mindedness: The Inquiring Mind*. Just as in the earlier research on personality, this factor was defined by the smallest number of socio-emotional skill items, even though theoretical writings on 21st Century skills greatly emphasize the importance of intellectual curiosity and exploration as well as innovation and creativity. Nonetheless, the items classified in the *Innovation* facet here in Table 4.3 included interesting and novel features, such as having vision and insight, tinkering and learning from mistakes, and the excitement of creating something new. Interestingly, there were few, if any, of the usual Openness items that relate to the standard Openness facet of aesthetic interests or sensitivity, with its emphasis on art, music, and literature. It is possible that these characteristics are underrepresented in more technology and employment oriented collections of 21st century skills; instead, this facet emphasized appreciating beauty, living in harmony with nature, and emotions relevant to spirituality, such as awe, wonder, and reverence. Also of interest is an additional facet that we have included in the present OECD framework, namely self-reflection and awareness of self and inner experiences (like the facet *Openness to feelings* that is included on the NEO PI-R). We selected that facet because it also appeared in our review of the literature on emotional competence (see below).

In conclusion, these broad domains defined by socio-emotional skill characteristics bear enough similarity to the familiar and well-studied Big Five personality to give us some confidence about their likely replication and generalizability. At the same time, the content of these five socio-emotional skill factors emphasizes their unique origin in 21st century skills and positive-psychology, with its approach based in strengths and virtues (Seligman, Steen, Park, and Peterson, 2005), and can thus advance our understanding beyond from the hierarchical personality taxonomy of the Big Five and the three core facets shown in Tables 4.1 and 4.2. The socio-emotional characteristics summarized in Table 4.3 provide a starting place for a new integrative and operational definition of socio-emotional characteristics that can be implemented in a longitudinal study on course and impact of social-emotional skills. More generally, socio-emotional skills are best defined as *individual characteristics that (a) originate in the reciprocal interaction between biological predispositions and environmental factors, (b) are manifested in consistent patterns of thoughts, feelings and behaviors, (c) continue to develop through formal and informal learning experiences, and (d) influence important socioeconomic outcomes throughout the individual's life* (OECD, 2015; De Fruyt, Wille, and John, 2015; Primi, Santos, John, and De Fruyt, submitted).

Reasons for Beginning the Conceptual Framework with the Big Five Facets as a Minimal Set

Previous OECD reports (e.g., Kautz et al., 2014) as well as economists working on socio-emotional skills have simply adopted the Big Five personality dimensions as the conceptual framework of choice because they found that most of the empirical research on the development and longer-term impact of socio-emotional characteristics has been conducted with Big Five measures. Specifically, in Kautz et al.'s (2014) report *Fostering Non-Cognitive Skills to Promote –Lifetime Success*, five economists from three countries concluded:

“Although non-cognitive skills are overlooked in most contemporary policy discussions and in economic models of choice behavior, personality psychologists have studied these skills for the *past century*.

Psychologists primarily measure non-cognitive skills by using self-reported surveys or observer reports.

They have arrived at a relatively well-accepted taxonomy of non-cognitive skills called the Big Five, with the acronym **OCEAN**, which stands for: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism.”

Similarly, the recent report by the National Academy of Sciences (2012) in the United States, entitled “*Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*”

observes: “For the past two decades, the big five model of personality has been widely accepted as a way to characterize competencies in the interpersonal and intrapersonal domains” (p. 28).

In contrast to the view, commonly held by laypersons, that personality traits are fixed, this report further cites extensive research demonstrating that personality traits are malleable and subject to multiple influences even during adulthood (e.g., Srivastava et al., 2003), and concludes that “these traits can be altered by experience, education, parental investments, and targeted interventions” (p. 24). Reviewing the available research evidence, the report then concludes:

“The five major factors provided a *small number of research-based constructs* (emphasis added) onto which various terms for 21st century skills could be mapped. The facets helped to define the range of skills and behaviors encompassed within each major factor to serve as a point of comparison with the various 21st century skills.”

Further, important evidence comes from the pilot study conducted jointly by the OECD and the Ayrton Senna Institute in Brazil, which sampled more than 21,000 students from 5th to 12th grade in public schools in the State of Rio de Janeiro (OECD, forthcoming). This research used both careful conceptual reviews and extensive analyses of the new data collected in Brazil. It reached a similar conclusion, namely that the Big Five, supplemented with an assessment of positive core self-evaluations (e.g., self-esteem), would provide a comprehensive framework for the socio-emotional skills of children aged 10 to 19 in Brazil. A related background study (Primi and Santos, 2015) carefully reviewed all existing child and adolescent instruments assessing socio-emotional constructs and then selected the most promising ones, which were then translated into Portuguese. Through a series of pilot studies, these researchers arrived at a final set of 65 socio-emotional skills items for younger children (ages 9-14) and 96 items for older children (14-19). Remarkably, in their highly diverse samples of Brazilian children in public schools, analyses of these items again produced the now familiar Big Five, plus a 6th dimension capturing extremely negative core self-evaluations (i.e., low self-esteem and external locus of control). For example, their careful empirical mapping studies (see also Primi et al., 2014, submitted) showed that the scales on such child assessment instruments as the *Strengths and Difficulties Questionnaire* (SDQ) and the three-dimensional *Self-Efficacy Scales for Children* (SES) could be represented well within the more comprehensive Big Five model. Social skills items like “I am able to tell a joke or story to a group of friends at school” loaded with other items from the Engagement with others (or extraversion) factor; The Prosocial strengths (respect) item “I try to be nice to other people. I care about their feelings” loaded on the Collaboration (or agreeableness) factor. The time management item “I am able to complete all my homework” loaded with other items from the Task performance (or conscientiousness) factor; and the item indicating a lack of emotion regulation skills “I get very angry and lose my temper” loaded on the Emotion regulation (or stability) factor.

In short, both conceptual and empirical evidence point to the promise of a framework that has an empirical foundation in the insights and 20-year research record accumulated for the Big Five. However, it is the view of the expert committee charged with developing the conceptual framework for the new OECD longitudinal study of skill development that this framework ought to *go beyond the Big Five*, in two ways. First, in contrast to the two previous OECD reports (Kautz et.al, 2014; OECD, 2015b), which examined only the superordinate-factor level of the Big Five, we propose to assess socio-emotional skills not only at the factor level but also at the lower level of more specific facets (as illustrated in Tables 4.2 and 4.3). Second, in the remainder of this report, we will conduct comprehensiveness checks of the proposed framework by reviewing other frameworks, approaches, cultural perspectives, and research findings, supplementing the initial draft framework based on the Big Five with additional constructs as needed. Before we turn to these comprehensiveness checks, we first review the evidence that is already available for the consequential validity of some of the socio-emotional skill concepts in the proposed framework.

Cross-cultural relevance of the framework

The OECD Longitudinal Study of Social and Emotional skills in Cities intends to launch longitudinal surveys in a variety of countries across the world. This multi-national approach poses specific challenges for the cross-cultural validity of the proposed framework and its constructs, and the psychometric requirements for its operationalization.

Is the proposed framework cross-culturally relevant?

Nowadays, there is evidence that the Big Five dimensions can be recovered in the cultural and language communities to be included in the OECD Longitudinal Study of Social and Emotional skills in Cities (e.g., Mexico, Brazil, South Korea, Japan, Norway, and Russia) and that in these countries the Big Five capture core qualities underlying a broad range of personality characteristics. Conceptually, the Big Five have to be understood as the basic colours of character and personality, expressed as more pure or blended manifestations of individual differences at the phenotypic level. Although the Big Five are a guiding framework to structure personality descriptions across the world, this does not necessarily imply that these are the only important constructs to consider.

Reviewing the literature, it is clear that, when transported to another culture, the Big Five factors can be recovered in self and observer ratings. Early studies compared the factor structure of personality ratings across individual countries, such as the USA, Spain, and Mexico (e.g., Benet-Martinez and John, 1998). More recently, large international teams of investigators have collaborated on larger-scale studies. Schmitt and colleagues (2007), for example, could retrieve the five-factor structure when the Big Five Inventory (BFI) was translated from English into 28 different languages administered to 17,837 individuals from 56 countries, including Argentina, Australia, Austria, Bangladesh, Belgium, Bolivia, Botswana, Brazil, Canada, Chile, Croatia, Cyprus, Czech Republic, Democratic Republic of Congo, Estonia, Ethiopia, Fiji, Finland, France, Germany, Greece, Hong Kong, India, Indonesia, Israel, Italy, Japan, Jordan, Latvia, Lebanon, Lithuania, Malaysia, Malta, Mexico, Morocco, Netherlands, New Zealand, Peru, Philippines, Poland, Portugal, Romania, Serbia, Slovakia, South Africa, South Korea, Spain, Switzerland, Taiwan, Tanzania, Turkey, Ukraine, United Kingdom, United States and Zimbabwe. Likewise, McCrae and colleagues (2005) found support for the FFM structure when analysing college-students' NEO-PI-R descriptions (N=11,985) of college-aged (18-21) and adult-aged (>40 years) individuals from 50 nations, including Argentina, Australia, Austria, Belgium, Botswana, Brazil, Burkina Faso, Canada, Chile, China, Croatia, Czech Republic, Denmark, Estonia, Ethiopia, France, Germany, Hong Kong, Iceland, India, Indonesia, Italy, Japan, Kuwait, Lebanon, Malaysia, Malta, Mexico, Morocco, New Zealand, Nigeria, Peru, Philippines, Poland, Portugal, Puerto Rico, Russia, Serbia, Slovakia, Slovenia, South Korea, Spain, Switzerland, Thailand, Turkey, Uganda, United Kingdom: England, United Kingdom: Northern Ireland, and the United States. De Fruyt et al. (2009) provided similar evidence analysing descriptions of adolescents (12-17) obtained in 24 cultures, including Argentina, Australia, Chile, People's Republic of China, Croatia, Czech Republic, Estonia, France, Hong Kong, Islamic Republic of Iran, Japan, Malaysia, Peru, Poland, Portugal, Puerto Rico, Russia, Serbia, Slovak Republic, South Korea, Thailand, Turkey, Uganda, and the USA. The explicit listing of countries illustrates that these are well-spread across North and South America, Western, Eastern and Southern Europe, the Middle East and Africa, Oceania, and South/South-East Asia and East Asia.

Whereas the previous work concentrated on the replication of the structure of imported inventories constructed in the US (i.e. the BFI or the NEO-PI-R) and administered to adolescents (De Fruyt et al., 2009) or adults (Schmitt et al., 2007; McCrae et al., 2005) in a broad variety of cultures, the International Consortium for the Developmental Antecedents of the Five-Factor Model (ICDA-FFM; Kohnstamm, Halverson, Mervielde, and Havill, 1998) worked bottom-up in various cultures examining the content and structure of parental free descriptions to define the structure of personality for children. This group of

developmental, temperament and personality researchers examined the active instead of the passive (like represented in dictionaries) personality descriptive vocabulary. At the same time, they aimed to assemble item sets that represented age-specific indices of individual differences that were more sensitive to describe developmental differences in childhood. Nearly all studies conducted before 1995 had used adjective lists or inventories initially developed for adults preventing the emergence of childhood specific personality dimensions or facets.

Their methodology was simply asking parents with children in the age-range of 6 to 12 (primary school) to describe what they found characteristic of their child, without any additional prompts to avoid influencing the content of their descriptions. All descriptors collected this way were subsequently sorted in a personality descriptive lexicon including 14 major categories, containing the Big Five supplemented with a number of temperament categories. This age-grouped descriptor database served as the starting point for developing age-specific item sets to enable bottom-up research on the structure of personality in youth. The ICDA-FFM group used this approach in various countries, including Belgium, China, Germany, Greece, the Netherlands, Poland, and the United States, to be in a position to study the impact of culture on the resulting structure. This basic and innovative approach provided a strong test of the validity of the comprehensiveness of the Five-Factor Model from a cross-cultural *and* a developmental perspective.

Mervielde and De Fruyt (1999, 2002) used this approach and assembled a pool of near to 10.000 parental free descriptions of Flemish children aged between 6 and 13 years. Lexicon categories were further split in about 100 homogeneously descriptive categories, and their content was represented by 2 to 3 personality items for each age-group (6 year-, 9-year, and 12-year olds). These item sets were subsequently administered to large samples of parents and teachers requested to rate children aged 6 to 12 years. An analysis of the factor structures within and across age groups and gender clearly pointed towards the same five factors, identified as extraversion, benevolence (agreeableness), conscientiousness, emotional stability or neuroticism, and imagination (openness). Benevolence referred to a broader set of traits than the adult agreeableness factor, referring to content associated with the concepts 'easy-difficult' child described in the temperament literature (Thomas, Chess, Birch, Herzig, and Korn, 1963) and 'manageability' from the perspective of the parent or teacher informant.

Despite the initial focus to work with trait indicators for specific age-groups, the resulting sets represented a very similar behavioural content and a highly similar higher-order structure. Mervielde and De Fruyt (1999, 2002) additionally examined the lower-level structure across age-groups and proposed a common set of 18 facets (with 8 items per facet) to describe children from 6 to 12 in the Hierarchical Personality Inventory for Children (HiPIC; Mervielde and De Fruyt, 1999; Mervielde, De Fruyt, and De Clercq, 2009). The emotional stability domain included two facets, i.e. anxiety and self-confidence, whereas extraversion grouped four facets: energy, expressiveness, optimism and shyness. Imagination included creativity, intellect and curiosity, whereas Benevolence distinguished among altruism, dominance, egocentrism, compliance and irritability. Finally, conscientiousness grouped the facets concentration, perseverance, orderliness and achievement striving. The American team of the ICDA-FFM, led by Halverson, constructed the Inventory for Child Individual Differences (ICID; Halverson et al., 2003), following a similar starting point. The resulting factor solution of the ICID was comparable to that of the HiPIC (Tackett, Kushner, De Fruyt, and Mervielde, 2013).

When researchers operationalized personality with indigenous concepts rooted in a particular culture, some studies have encountered difficulties recovering all of the Big Five factors, or have found factors that were slightly different from those in Western studies. Cheung and colleagues (2001), for example, jointly examined the structure of the NEO-PI-R and the Chinese Personality Assessment Inventory (CPAI; Cheung et al., 1996). Their results provided powerful evidence for the replication of four of the Western Big Five domain (namely, Neuroticism, Conscientiousness, Agreeableness and Extraversion). In addition their results suggested a separate 'interpersonal relatedness' factor (see the Relationship Harmony concept

included in the present framework) and three specific Openness facets (O3: Feelings, O2: Aesthetics, and O1: Fantasy) chiefly loading an openness factor. The additional factor interpreted as “interpersonal relatedness” had loadings only from CPAI scales, including Harmony, Optimism (versus Pessimism), Ren Qing (relationship orientation), Flexibility, Defensiveness (Ah-Q mentality), Face, and Logical versus Affective Orientation. This additional factor could not be absorbed by the NEO-FFI factors across three different samples; the variances explained by the FFM ranged from as low as .08 (Ren Qing; sample of 372 Chinese managers) to .31 (Flexibility; same sample).

A project with impressive bottom-up work has been conducted in South Africa, making use of the lexical approach to identify the local vocabularies of personality description and then derive basic dimensions of personality from that emic material. Nel et al. (2012) developed the South African Personality Inventory (SAPI) by starting with the personality descriptive language obtained in semi-structured interviews of more than 1,200 individuals representing the 11 major language groups in South-Africa. This vocabulary was grouped into 9 broad content-based clusters: conscientiousness, emotional stability, extraversion, facilitating, integrity, intellect, openness, relationship harmony, and soft-heartedness. Those clusters that were not represented by the Big Five were all related to social-relational functioning and tapped into ways of maintaining positive interpersonal relationships with others, and could thus be conceptually linked to the Agreeableness in the Big Five.

Valchev and colleagues (2014) conducted a series of follow-up studies, and reported that the social-relational scales of the SAPI generated two factors not presented in the Big Five, assessed by the BTI (Taylor and De Bruin, 2005) with items from the International Personality Item Pool (IPIP; Goldberg et al., 2006). The positive social-relational factor was defined by facilitating, integrity, relationship harmony, active support, empathy, facilitating, and integrity, and South-African Blacks scored higher on this factor than did Whites. Moreover, social-relational concepts explained substantial variance in pro-sociality beyond the Big Five. Finally, there was evidence that the Big-Five-Plus-Two factor-structure could be also recovered from a joint factor analysis of 50 IPIP items and SAPI social relational scales administered to a mixed sample of 452 mainstream Dutch, 427 Western, 225 Antillean, Surinamese, and Indonesian, and 179 non-Western participants. In sum, evidence for the cross-cultural generalizability of the basic Big Five taxonomy is substantial overall, but *additional dimensions may be necessary to provide a comprehensive structural representation of personality in particular cultures*. The studies reviewed so far suggest the most promising candidates for additional constructs are likely to be found in the relational domain.

More specifically, psychological studies of cultures suggest that the structure and meaning of social relationships may differ across cultures (e.g., Markus and Kitayama, 1991; see Benet-Martinez and Oishi, 2008 for a review). We therefore expect that additional constructs may be needed particularly in the interpersonal domain, supplementing constructs broadly situated in the Agreeableness domain. One important candidate construct involves the *construal of self as interdependent in relationships* (e.g., Markus and Kitayama, 1991; Singelis, 1994); research has shown that individuals in East Asian societies with more “collectivistic” values (e.g., China, Korea, Japan) score higher on interdependence than individuals in Western societies with more “individualistic” values (e.g., USA; Australia); the opposite pattern is obtained for interdependent self-construal. These variables are potentially important because they can help explain, at the individual level, important cultural differences in life satisfaction and well-being, which are important life outcome indicators and considered central by the OECD.

Another promising candidate is *relationship harmony*, mentioned above in the South African studies. Kwan, Singelis, and Bond (1997) proposed a relationship-based measure for this concept and compared East Asian countries (e.g., Hong Kong) with the US. Their careful analyses showed that self-esteem influenced life satisfaction more than relationship harmony in the US whereas both factors were equally important in Hong Kong (Kwan et al., 1997). They also measured individual differences in independent and interdependent self-construal in both cultures; the effect of *independent* self-construal on life

satisfaction was mediated through *self-esteem*, whereas the effect of *interdependent* self-construal was mediated through relationship harmony. In short, these variables capture cultural differences of “self-in-relationships” that can be measured at the level of the individual. Indeed, members of East Asian cultures, even when they live abroad in individualistic cultures (like Asian-Americans in the US) consistently report somewhat lower levels of life satisfaction, happiness, and self-esteem than do Western countries (like the US and Western Europe), and these kinds of differences need to be anticipated and conceptualized.

More recently, one additional concept has been suggested in the *Journal of Cross-Cultural Psychology* and there is reason to believe it will be relevant to the Emotion Regulation domain in the present framework. Specifically, it is *fear of happiness*, which can be measured with a short self-report scale (Joshanloo and Lepshokava, 2014). This research on individual differences is consistent with the work of Jeanne Tsai (e.g., Tsai, Knutson, and Fung, 2006), who suggested that cultures differ in the way they value particular emotions, especially emotions considered positive and exciting (or arousing) in the West, especially intense happiness states like joy, enthusiasm, and excitement (as well as love and pride). East Asian cultures have been shown to value these intense positive emotions less than Western cultures and thus express them less in publically observable behaviour. The Fear of Happiness scale explicitly measures various specific beliefs about happiness that individuals may learn, to varying degrees, from their culture and socialization experience. Some individuals tend to be more suspicious of feelings of great happiness, expecting something bad to happen when they allow themselves to be too happy, whereas others tend to embrace and indulge feelings of intense happiness. As expected, in an initial cultural comparison, East Asian countries (e.g., Japan) scored substantially higher on fear of happiness than Westerners (e.g., Western Europe), who in turn scored higher than Brazilians who had by far the lowest scores of all countries studied and seemed to embrace happiness without fear or suspicion (Joshanloo and Lepshokava, 2014). These cultural differences were substantial in size but need to be considered with caution until replicated. However, they hold the promise to understand, at the level of the individual, why East Asian consistently report lower levels of positive emotion and many other positively balanced? attributes, such as self-esteem, life satisfaction, well-being, and even extraversion.

It is important to realize that this research on interdependent self-construal and relationship harmony did *not* suggest that these constructs could not be measured reliably in one of the two cultures involved. On the contrary, reliable measurement in both cultures made possible the demonstration of mean-level cultural differences and their explanations. In other words, individual differences in interdependence and relationship harmony exist and can be measured in the USA but they are less important and therefore less expressed there than in China. Indeed, the discussion on structural replicability (or invariance) has been complicated by arguments about the importance of particular factors across cultures. Structural replicability does not imply that factors have equal importance across cultures. For example, more collectivistic cultures may put higher value on Agreeableness facets related to politeness and compliance, whereas more individualistic cultures may value individual achievements and thus put higher value on traits associated with “standing out” from the group, such as assertiveness (i.e., Extraversion) and achievement striving (i.e., Conscientiousness). These weighting differences reflect important cultural variations but the Big Five may be structurally replicable in both groups of cultures.

Measurement invariance and comparisons between- and within- cultures

Cultural differences not only play a role when comparing scores of individuals between cultures and countries, but societies within single countries have also become increasingly heterogeneous the past years in terms of the cultural backgrounds of their members. Half of the population at schools in capital cities, for example, may have various ethnic origins and pupils’ cultural identity may be a mixture of characteristics from the host and the culture of origin. Moreover, people within a culture, may identify with multiple groups at the same time, such as ethnic origin, ethnic identity, gender, sexual orientation, age, and social-economic group. These different group attributes may interact and affect individuals’ score patterns

on psychological constructs (De Fruyt and Wille, 2013), introducing different forms of construct, method and item biases (Van de Vijver and Leung, 1997). The absence of bias is labelled as equivalence or measurement invariance.

Construct bias refers to the phenomenon that constructs only partly share meaning across groups. For example, the trait of Assertiveness is a relatively factor pure indicator of Extraversion in US samples, though taps into Extraversion and Emotional stability in Germanic languages like Dutch and German. Method biases refer to the differential impact on groups of the scale formats that are used or the way the assessment is conducted. For example, some groups may find it difficult to use a sorting procedure like the Q-sort, sorting different values, or choosing the item “most like you” and “least like you” from an ipsative item set. A final threat to measurement equivalence is differential item functioning (DIF): “DIF occurs when individuals with the same level or amount of a trait, but from different cultural groups, exhibit a different probability of answering the item in the keyed direction” (Church, 2010, p. 154).

Measurement invariance can further be demonstrated at different levels (Vandenberg and Lance, 2000), distinguishing among configural (same number of factors and pattern of loadings), metric (loading patterns constrained to be equal across groups), and scalar invariance (item intercepts are equal across groups). Church and colleagues (2011) recently examined DIF in the data that were collected in the USA, the Philippines, and Mexico, and showed that 40 to 50 percent of NEO-PI-R (Costa and McCrae, 1992) items exhibited some form of DIF, suggesting that one should be careful with making comparisons of mean scores across cultures. For example, Schmitt et al. (2007) compared self-reported BFI means observed in 10 world regions and found that the level of negative affect was very high in Japan, but also that East-Asians scored the lowest on conscientiousness, whereas the mean for Africans was the highest. Schmitt et al. (2007) raised that it is unlikely that Japanese would be perceived as low in self-discipline, order and achievement, and suggested that for some constructs and in some cultures, culturally endorsed response styles may be responsible for such effects. There are alternative, powerful methods to examine measurement invariance across cultural groups, such as the bilingual approach where the same samples are administered the same instrument in each of the two languages (e.g., see Benet-Martinez and John, 1998, Studies 2 and 3). When assessments are equivalent, the resulting mean scores should be alike (McCrae and Terracciano, 2008). More recently, Kyllonen and Bertling (2013) have suggested the use of vignettes to correct for potential biases in the use of response scales across cultures, and Primi et al. (2014, under review) have demonstrated that the vignette approach can be effective in studies of personality self-ratings made by children.

4.2. High Predictive Power and Comprehensiveness

Although the description of social-emotional skill development trajectories is a valuable research objective in itself, the OECD’s Longitudinal Study of Social and Emotional Skills in Cities also aims to clarify the mechanisms of how these trajectories lead or contribute to a broad series of consequential outcomes. National contexts, in addition to individual circumstances, are considered in this respect as moderators and mediators of the outcomes.

Social-emotional skills may affect the listed outcomes below in multiple ways at the same time, but the driving mechanisms of outcomes can also differ across development. Social-emotional skills may affect outcomes directly or influence an outcome indirectly via another construct. Such effects can be independent effects, i.e. one or two social-emotional skills exerting an influence independent of each other, but there may be also interactive effects, substituting (resource substitution model) or strengthening (Matthew) effects. Social-emotional skills may further act as a moderator of an association between a predictor and outcome. Traits and skills will not only affect the outcomes, but will also be shaped by the outcomes (i.e. reciprocal relationships). Finally, social-emotional skills can have short and long term effects.

Evidence on the predictive power of social and emotional skills

Educational attainment

School systems around the world focus on both knowledge acquisition and developing social-emotional skills in youth. In the impressive meta-analysis summarizing results of 213 school-based social-emotional learning programs conducted with pupils (N= 270,034) from kindergarten to high school already discussed in the previous section, Durlak and colleagues (2011) reported the effects of social-emotional learning ranging from .22 (conduct problems) to .57 (social-emotional skills). They also found, however, that training of social-emotional skills had direct effects on academic performance. Social-emotional skills are hence both targets and means in formal education programs.

There is a growing interest in educational psychology to examine more social-emotional variables, including interests, motivational factors and personality traits to explain the trends/outcomes? in educational attainment. This research, focusing more in the typical behavioural indices affecting the study performance, complemented the well-established research line on more maximal predictors of learning outcomes such as cognitive abilities. Research by Strenze (2007), for example, showed that intelligence explained nearly a quarter of the academic attainment scores, with a corrected correlation coefficient of .56. A key question was to what extent does academic achievement also predicted by socio-emotional skills.

Although a series of specific traits have been examined with respect to educational outcomes, such as procrastination (Steel, 2007), grit (Duckworth and Seligman, 2006; Duckworth, Peterson, Matthews, and Kelly, 2007), and goal setting-engagement (Bipp and Van Dam, 2014), most work has been conducted using the broad Big Five domains rather than the specific facets, with a large number of replicated studies now dating back to more than 20 years (e.g., John et al., 1994). Poropat (2009) meta-analytically investigated the relationship between the dimensions of the FFM and academic performance. Some of the primary studies included in this meta-analytic summary also reported correlations between intelligence and academic performance, and were also quantitatively summarized. The sample-weighted correlation corrected for scale reliability between intelligence and academic performance was .25, hence substantially lower than the .56 reported by Strenze (2007). The larger effect size estimated by Strenze may be due to sample variation or to the absence of range restriction, so the .25 reported by Poropat for intelligence provides a benchmark to interpret the relative weight of the FFM traits. The corrected correlations for the FFM scales and academic performance were .22 for Conscientiousness, .12 for Openness, .07 for Agreeableness, .02 for Emotional Stability, and -.01 for Extraversion; three of the Big Five, namely Conscientiousness, Openness, and Agreeableness were significant explanatory constructs. The correlation found for Conscientiousness alone almost equalled the one found for intelligence. Academic level was found to significantly moderate the FFM-academic achievement association, with the largest coefficients found in primary education for both intelligence and all FFM factors, and declines from primary to secondary and tertiary level for Intelligence, Agreeableness, Emotional stability, and Extraversion, and linear declines across the three levels for Openness. The correlation for Conscientiousness did not significantly alter across academic levels. The correlations in primary education were .58, .30, .28, .20, .18, .24 between academic achievement and respectively intelligence, Agreeableness, Conscientiousness, Emotional Stability, Extraversion and Openness. Poropat (2009) further found that Conscientiousness added little to the prediction of tertiary GPA when partialled out for secondary GPA, though still slightly performed better than intelligence. In a meta-analytic investigation of adult-rated child personality and academic performance in primary education (Poropat, 2014a), corrected correlations of .43, .18 and .50 were reported for Openness, Emotional Stability and Conscientiousness respectively, significantly outperforming the effects observed for self-ratings in the case of Openness and Conscientiousness. These relationships were not moderated by age or year of education (grades 1 to 7). The positive associations

between Conscientiousness and Openness on the one hand and academic achievement on the other were also extended to other-rated personality (Poropat, 2014b).

A broad series of studies have directly related personality traits to education performance outcomes, with the FFM Conscientiousness, Openness and Neuroticism factors as key dimensions to describe achievement-relevant personality (Briley, Domiteaux, and Tucker-Drob, 2014). Spengler and colleagues (Spengler, Ludtke, Martin, and Brunner, 2013), following a large representative sample of 15-year-old students and another sample of students of the 9th and 10th grade, showed that conscientiousness better predicted grades ($r = .15-.30$), whereas openness was more strongly associated with performance (.15-.32) on math and reading items culled from the Programme for International Student Assessment (PISA; OECD, 2009).

Performance at work and employability

Industrial and Organizational psychologists have been interested for a long time in the predictors of job performance when assigning job applicants to vacancies. More recently, due to the volatile and changing employment market and requirements of life-long learning, this professional group got also strongly interested in the concept of ‘employability’, referring to individuals’ labour market fitness and requirement to be in charge of their own careers. The interests of industrial and organizational psychologists hence overlapped considerably and increasingly with the concept of 21st century skills, given their increased attention for meaning of working (work or employment flows better?) for people, happiness and mental health at work, and work-life balance, going beyond mere indicators of job performance such as quantity and quality of task performance.

Parallel to the educational domain, also the Industrial and Organizational field do not have an agreed upon taxonomy to describe the basic qualities underlying individual difference predictors of job performance and employability. Selection psychologists and human resources professionals frequently introduce a new vocabulary to refer to the individual qualities they are looking for, labelling these the past 10 years as (behavioural) ‘competencies’, and more recently as ‘talents’ but also as ‘21st century skills’. Although the introduction of these new concepts underscores the dynamics of this field, it does not necessarily facilitate the communication among various labour market stakeholders, and it imposes major challenges for assessing these qualities reliably and efficiently among job applicants or incumbents. There is usually not a one-to-one relationship between the listed competencies, socio-emotional skills, or talents on the one hand and the concepts and constructs for which differential psychologists have constructed particular models and assessment tools. Competencies like “Resiliency or low stress Vulnerability”, for example, are very trait-like and easy to map into a personality descriptive model such as the FFM, and thus relatively straightforward to assess. However, competencies such as “having Impact”, “having a Helicopter view” or ability to articulate “Vision” or “Deal with disputes” are more complex “hybrid” constructs that are best conceived as blends between more cognitive and more trait-liked characteristics. These latter examples nicely illustrate that the bifurcation between cognitive and non-cognitive factors is artificial, and often rather useless. Skills that help turn a conflict into a fruitful discussion require abilities to analyze, to communicate, and also to regulate emotions, hence, requiring a mixture of cognitive and non-cognitive resources in a single labelled skill.

Selection and assessment psychologists, with a background in individual differences, started to develop conceptual models to relate competency models to constructs from differential psychology. Hoekstra and Van Sluijs (2003) consider FFM personality traits and intelligence as building blocks of behavioural competencies, with formal and informal learning processes impacting upon the competency level during development. De Fruyt, Bockstaele, Taris and Van Hiel (2006) illustrated how this model can be used to link the FFM trait model and police interview competencies. The central position as building blocks of competencies given to traits and intelligence in Hoekstra and Van Sluijs’ model (2003) makes

these constructs important assets to examine an individual's employability. These approaches further illustrate that the applied field can borrow constructs and assessment methodology from the differential psychology fields.

Nowadays, there is convincing evidence in Industrial and Organizational Psychology that individual differences' constructs are key variables to assess when discussing performance at work and employability. There is considerable meta-analytic evidence that cognitive abilities are among the best predictors of job performance and training proficiency across a range of jobs and in different cultures (Salgado, Anderson, Moscoso, Bertua, and De Fruyt, 2003; Salgado et al., 2003), with also good validity early on to predict outcomes that are observed later in life such as income and occupational attainment (Judge, Higgins, Thoresen, and Barrick, 1999; Woods, Lievens, De Fruyt, and Wille, 2013). Validities for intelligence measures and job performance do not vary much across jobs (Salgado et al., 2003), showing that intelligence is predictive across jobs, though type of job moderates this relationship, with stronger associations for more complex occupations. In addition to explaining job performance directly, there is also evidence that intelligence is probably more important when individuals get a new employment, because cognitive capacity facilitates learning and helps the individual adapt to new challenges, whereas its validity seems to decline somewhat when individuals get more acquainted with their job, underscoring the importance of cognitive ability for labour market fitness or employability.

In a meta-analysis of FFM measures and job performance, Hurtz and Donovan (2000) reported estimated mean operational (true) validity coefficients (corrected for unreliability in the criterion measure and range restriction on the personality measures) for self-reported personality of .22 (Conscientiousness), .14 (Emotional Stability), .10 (Agreeableness), .09 (Extraversion), and .05 (Openness). However, using theory to align Big Five dimensions to specific job-performance aspects, Hogan and Holland (2003) reported true estimated validities of .43 (Emotional stability), .35 (Extraversion-Ambition), .34 (Agreeableness), .36 (Conscientiousness) and .34 (Intellect-Openness to experience). These findings unequivocally demonstrate that the underlying broad dimensions of personality description are related to various performance indicators valued by employers.

In the past years, there has been increasing efforts to raise the predictive validities of personality measures to understand performance at work. First, there was considerable debate about the level of the trait hierarchy at which predictions of job performance were best made. That is, at the level of broader domain factors such as the Big Five dimensions or at the level of more specific facets. A similar discussion was conducted at the level of the outcomes, further subdividing job performance into quantity and quality of task performance, in addition to contextual and adaptive performance. The latter approach is also in line with Hurtz and Donovan (2000) arguing to use theory to better align the predictors and outcomes.

A second improvement involve contextualized personality assessments; that is, rather than asking for a description of one's personality "overall" or "in general," the more specific and relevant *work context* is included as the frame-of-reference for the personality description (Lievens, De Corte, and Schollaert, 2008), for example by adding a tag "at work" to items. Not surprisingly, these more contextualized personality assessments are better aligned with the work criteria they are supposed to predict and thus show generally better predictive validities. De Fruyt and Rolland (2013) illustrated the combined effects of aligning predictors and criteria and using a work frame-of-reference, showing that self-rated Conscientiousness at work correlated .36 with colleague-rated task performance (relative to a correlation of .27 using a general personality Conscientiousness scale). Self-rated Neuroticism and Openness to experience correlated -.21 and .26 with adaptive performance rated by colleagues (relative to correlations of .16 and .12, respectively, for non-contextualized general measures). A third established improvement is to include additional observers beyond self-descriptions (Connelly and Ones, 2010). Oh, Wang and Mount (2010) meta-analytically examined the effect of adding 1 to 3 observer personality ratings to self-ratings, convincingly showing that validity increases adding observers, reporting coefficients of .41

(Conscientiousness), .24 (Emotional stability), .34 (Agreeableness), .29 (Extraversion) and .29 (Openness/Intellect) when 3 observer ratings were added to the earlier reported self-ratings by Hurtz and Donovan (2000). Oh and colleagues' meta-analysis hence suggest that all FFM traits explain job performance ratings, and should be incorporated in a comprehensive assessment of a job candidate to evaluate employability.

Whereas the previous innovations are getting progressively integrated into professional assessment practice, a few other routes are currently under investigation, and their merits to improve prediction have yet to be established. A primary group of innovations is situated at the assessment side, examining the predictive validity of implicit measures of personality (Vecchione, Dentale, Alessandri, and Barbaranelli, 2014) or using a situational judgment paradigm (Lievens and Sackett, 2012) to describe an individual's trait positions. These approaches try to find alternative indices of personality beyond the traditional questionnaire approach. A second valuable line of research is more state-oriented research expanding the traditional between-individual paradigm with approaches to look at within-individual variability (Debusscher, Hofmans, and De Fruyt, 2014; Minbashian and Luppino, 2014; Minbashian, Wood, and Beckmann, 2010). These research lines suggest that in addition to between-individual differences (i.e. some people are generally more conscientious than others), there is also a huge variability within the person (i.e. a person may have a certain variability in conscientiousness during the day).

De Fruyt, Wille and John (2015) defined employability in terms of five key characteristics, defined as whether the person (a) demonstrates task-engagement and goal-setting, (b) can get along with other people, (c) adapts to/fits in an organizational structure, or has the capacity to deploy such structure (for those pursuing in self-employment), (d) learns on the job and can prepare for future challenges, (e) can deal with short and long term perspectives. This minimal set of employability indices taps into all basic personality dimensions, with task-engagement and goal-setting related to Conscientiousness, interpersonal skills related to emotion regulation and the core dimensions of the interpersonal circumflex (Extraversion and Agreeableness), and 'fitting in', 'learning and adapting', and 'time perspective' related to Openness to experience and Conscientiousness. Employability and the impact of FFM traits is further reflected in how individuals navigate on the employment market and develop their career paths (Wille, De Fruyt, and Feys, 2010, 2013).

Finally, two other prominent related though distinct areas indexing employability are the demonstration of entrepreneurship or self-employment. Again, the FFM dimensions turned out to be associated with these outcomes (Obschonka, Schmitt-Rodermund, Silbereisen, Gosling, and Potter, 2013; Obschonka, Schmitt-Rodermund, and Terracciano, 2014; Obschonka, Silbereisen, and Schmitt-Rodermund, 2012).

Mental health

There is strong evidence that social-emotional skills and their underlying trait building blocks show strong relations to a variety of mental health problems. Tackett (2006) and De Bolle, Beyers, De Clercq, and De Fruyt (2012) described five different ways how basic personality dimensions are connected to broad psychopathology dimensions, such as internalizing and externalizing behaviour, but also to specific disorders. The *vulnerability* model stipulates that particular traits make an individual more susceptible to develop a particular form of disorder. For example, there is convincing evidence that those higher on neuroticism have an increased likelihood to develop one or more depressive episodes later in life (Fanous, Neale, Aggen, and Kendler, 2007; Kotov, Gamez, Schmidt, and Watson, 2010). In the Tracking Adolescents Individual Lives' Survey (TRAILS), Laceulle, Ormel, Vollebergh, van Aken, and Nederhof (2014) found that personality at age 11 was predictive of internalizing and externalizing disorders at age 19. They also showed that changes in personality between age 11 and 16 were predictive of both internalizing and externalizing psychopathology between 16 and 19, controlling for basal personality

scores, demonstrating that both cross-sectional and dynamic trait indicators are indicative of later forms of psychopathology.

The *pathoplasty* mechanism further explains how traits can affect the manifestation, course and prognosis of mental disorders, although both may have independent origins and developmental paths. A well-documented example is the co-occurrence of callousness-unemotional traits and the diagnosis of conduct disorder in adolescence, having a worse prognosis (Hawes, Price, and Dadds, 2014). The direction of the effect can also be the other way around, i.e. mental disorders affecting the underlying trait dimensions. This causal model is also called the *complication* or *scar* model, with the disorder leaving a ‘scar’ on the individual’s personality. For example, multiple depressive episodes may have their influence on the persons’ neuroticism score (Fanous, et al., 2007), or recurrent symptoms of paranoia may have affected a persons’ trust (agreeableness) level. A fourth mechanism proposes that traits and broad psychopathology dimensions show systematic phenotypic co-variation and may form a single continuum, i.e. the so-called *continuity* hypothesis. Evidence for such relationships between dimensions assessed by the HiPIC and the CBCL internalizing and externalizing dimensions of psychopathology has been provided by De Bolle et al. (2012). Finally, an extension of this continuity hypothesis is the *spectrum* model, assuming that trait and disorder covariance has a common origin. This model has been recently supported in youth by Martell, Gremillion, Roberts, Zastrow, and Tackett (2014) describing longitudinal relations between personality traits and Attention-Deficit/Hyperactivity Disorder (ADHD) symptoms, and by Tackett and colleagues (2013) identifying common genetic influences on a general psychopathology factor and the negative emotionality trait in young twin pairs. This review of different mechanisms makes clear that the relationship between both sets of constructs is complex and De Bolle and colleagues (2012) convincingly demonstrated in youth that for many disorders, multiple of these mechanisms explain part of the association across time between traits and disorders.

Socio-emotional skills and their underlying traits also have many indirect effects, for example via the selection of specific situations that may independently or jointly contribute with skills/traits to mental health (De Fruyt and De Clercq, 2014). For example, an impulsive and aggressive adolescent may end up in a deviant peer-group, because s/he is rewarded and respected in this environment, further reinforcing dysregulation.

Physical fitness, health and longevity

In a recent review, Friedman and Kern (2014) distinguished among six core health outcomes covered in public health policy research: physical health, subjective well-being, social competence, productivity, cognitive function, and longevity. Currently there is substantial evidence that socio-emotional skills and their building blocks are associated with a broad range of health outcomes, such as smoking (Munafò, Zetteler, and Clark, 2007), obesity (Gerlach, Herpertz, and Loeber, in press; Sutin, Ferrucci, Zonderman, and Terracciano, 2011), alcohol craving and consumption (Papachristou et al., 2013; Stautz and Cooper, 2013), resilience to Alzheimer’s disease (Terracciano et al., 2013) and health status in cardiovascular populations (Versteeg, Spek, Pedersen, and Denollet, 2012). There is abundant evidence that traits related to impulsiveness are associated with eating problems and obesity. In a meta-analysis of 50 individual studies, Fischer, Smith, and Cyders (2008) found that emotion-based impulsivity indices such as negative urgency were substantially correlated with bulimic symptoms. Likewise, Sutin and colleagues (2011) followed 1988 individuals over a time span of more than 50 years and reported that low Agreeableness and high impulsivity traits predicted a larger increase in body mass index across adulthood.

Socio-emotional skills have been further successfully related to constructs from the positive psychology area, such as happiness (Hills and Argyle, 2001), quality of life, and subjective well-being (Diener, 2013; Steel, Schmidt, and Shultz, 2008). Many of these health and well-being related outcomes have subsequent effects on other criteria. For example, in a series of experiments, Oswald, Proto and Sgroi (2014) showed that happiness affects productivity.

Friedman and Kern (2014), however, warn against an oversimplified view of the association between dispositions and indicators of well-being and health, because many studies report the association between concurrently assessed self-reported dispositions and health indicators using the same informants and often using items that are included at both the “predictor” and the “outcome” side. This cautionary note aside, there is nevertheless enough evidence that dispositional constructs do contribute substantially to objective health indices. Friedman and Kern (2014) urge the field to consider more comprehensive and complex causal models of relationships among personality variables and correlated outcomes, taking into account mediator and moderator variables. For example, genetic predispositions, the environment and personality affect lifestyle patterns across time, manifested into correlated subjective well-being and physical health scores at Time 1 in development and these three may influence subjective well-being and physical health parameters observed at Time 2. Lifestyle patterns may mediate well-being/physical health from Time 1 to Time 2, and different contextual variables, such as psychological or biomedical intervention programs rolled out at school may moderate these developmental trajectories (Friedman and Kern, 2014).

In an impressive meta-analysis of 20 independent samples summarizing findings obtained on over 9000 participants, Kern and Friedman (2008) provided evidence on the lifelong significance of conscientiousness for individuals’ health and longevity. Results were straightforward and showed that higher levels of conscientiousness were significantly associated with longevity ($r = .11$, 95% confidence interval = .05-.17), with the strongest correlations observed for the goal achievement (persistent, industrious) and inhibitory (organized, disciplined) facets of conscientiousness. Kern and Friedman (2008) argue that the protective effects of conscientiousness may probably work via multiple ways across the life course ultimately contributing and combining into longevity. People high in conscientiousness for example may engage in healthier behaviours, or select safer and healthier family, living and work environments. Conscientiousness is a predictor of employability, and as a result persons with higher conscientiousness scores may end up in better jobs, having higher incomes and building more successful careers. Conscientiousness may further buffer and moderate the relationship between neuroticism and negative outcomes (e.g. poor mental health; see previous section). Another potential mechanism is that conscientiousness and health are influenced by similar genes, and hence are associated at the phenotypic level. Finally, and probably the most difficult to investigate, conscientiousness may contribute to health via the accumulation of small positive actions and/or the reduction of very small risks across the lifetime.

Civic engagement and environmental awareness

Civic engagement, environmental awareness and sustainable behaviour are outcomes that have become increasingly important during the past decade. Omoto, Snyder and Hackett (2010) examined motivational and personality predictors of activism and civic engagement, showing that other-focused motivation predicted AIDS activism and civic engagement better than self-focused motivation, interpersonal orientation and traits. Schnittker and Behrman (2012) examined the effects of schooling on civic engagement (participation and volunteering) and social cohesion (density of social network and quality of social relations) tempering somewhat previous optimism on the effects of education on achieving these outcomes. The effects of schooling on volunteering and participation in civic organizations disappeared almost entirely when taking into account different confounders. They concluded that increased schooling may generate some tension between navigating on the employment market and non-market commitments, as well as between independence and interpersonal reliability, making those who invest in schooling also more apt to pursue career-oriented interests, with less time left to engage in volunteering activities or civic engagement.

Developmental psychologists have paid attention to a related construct with high social significance called *generativity* (Erikson, 1950). During middle adulthood, somewhere between the ages of 40 and 65, people strive to create or nurture things that will outlast them. This can be achieved by having children or by contributing to positive changes that benefit other people, society in general, but especially future generations (e.g., building the Golden Gate Bridge). The generativity stage of development in

Erikson's model refers to "making your mark" on the world, through caring for others, creating things and undertaking things that make the world a better place. The lack of generativity, also described as *stagnation*, refers to failure of some individuals to find a way to contribute to these goals. These individuals may feel disconnected or disengaged with their community and even with the society as a whole (Van Hiel, Mervielde, and De Fruyt, 2006). Van Hiel and colleagues (2006) showed that the "making your mark" generativity construct was related to low Neuroticism (-.22), and high levels of Extraversion (.36); Openness (.21), and Conscientiousness (.26), but not to Agreeableness (.04).

Finally, raising environmental awareness and engagement has been promoted as one of the most recent major challenges to achieve in social-educational learning programs. Milfont and Sibley (2012) examined the relationships between Big Five traits and different indices of "green" (environmentally sound) behaviour at both the level of the individual and countries. At the level of the individual, they examined the association with valuing protecting the environment, whereas at the level of countries, they examined the association between aggregated personality traits (within countries) and country-level measures of sustainability, environmental attitudes, and values. At both levels of analyses, Agreeableness, Conscientiousness and Openness were significantly related to engagement in green behaviours.

Crime/safety

A final group of outcomes that are associated with social-emotional skills is the entire spectrum of externalizing disorders, including drug abuse, bullying, conduct problems, vandalism, youth and adult criminality, but also more socially camouflaged deviant behaviours such as unethical comportment, fraud, greed and corporate psychopathy (Furnham and Taylor, 2004). A key difficulty in this area of research has been the co-occurrence of symptoms and specific disorders making it very complex to study associations between deficiencies in specific skills and traits and particular disorders. For example, Attention-Deficit/Hyperactivity Disorder has shown high comorbidity with Oppositional-Defiant disorder, suggesting to structure childhood disruptive disorders along major dimensions (Martel, Gremillion, Roberts, von Eye, and Nigg, 2010). The common denominator across a broad range of studies (de Haan, Dekovic, van den Akker, Stoltz, and Prinzie, 2013; Decuyper et al., 2013; Decuyper, De Pauw, De Fruyt, De Bolle, and De Clercq, 2009; John et al., 1994; Klimstra, Luyckx, Hale, and Goossens, 2014; Nigg et al., 2002; van den Akker, Dekovic, and Prinzie, 2010; Van den Akker et al., 2013) is that across development, the externalizing spectrum is negatively related to Agreeableness and Conscientiousness and, depending on the type of disorder, positively with traits related to Neuroticism.

Distinctiveness from cognitive measures

The socio-emotional skill constructs should ideally exhibit incremental explanatory power over cognitive constructs. The Big Five dimensions can potentially fulfil this requirement, with the highest cross-correlation (between Openness and verbal IQ) being only about .30 (John et al., 1994; Loehlin, McCrae, Costa, & John, 1998). At the same time, the socio-emotional battery should include constructs that can be examined in interaction with cognitive measures to explain a variety of outcomes. For example, the interaction between basic cognitive ability and the socio-emotional differences in goal setting may help to explain school achievement, beyond the main effects of cognitive and socio-emotional measures (Poropat, 2009, 2014).

Learning from other conceptual frameworks

We review several other approaches to socio-emotional skills. Our goals in the following review will be to ensure that we identify any additional concepts with potentially high predictive validity at lower levels of abstraction, like those already presented in Figure 1.

Social-emotional Learning Approaches

We are now ready to return to the three models, mentioned earlier, that define social and emotional skills in terms of 5, 6, or even 8 domains. These three models are summarized in Table 4.4, each in one column. As shown there, Elias and colleagues (1997) described six major domains of social and emotional learning, defined as “core competencies to recognize and manage emotions, set and achieve positive goals, appreciate the perspectives of others, establish and maintain positive relationships, make responsible decisions, and handle interpersonal situations constructively.” More recently, in their review of learning effects in intervention programs, Durlak, Weisberg, Dymnicki, Taylor, and Schellinger (2011) explain that there is no single taxonomic or measurement model guiding past or present research on socio-emotional interventions; instead, interventions are typically driven by specific school or district contexts and needs, and thus tend to focus on a diverse set of outcomes that may vary as widely as obesity problems in elementary school, depressive symptoms in middle-school girls, or vandalism in high school classrooms.

Table 4.4. How Do the Major Domains of the Proposed Framework Relate to Models of Socio-Emotional Learning Domains and Emotional Competencies?

Socio-emotional framework	Socio-emotional learning models		Emotional competence models
	Elias et al. (1997)	Durlak et al. (2011)	Saarni (1999, 2011)
Engage-ment with Others(E)	4 - Establish and maintain positive relationships	4 - Relationship skills 3 - Social awareness	7 - Genuine emotion expression and reciprocity for relationships
Collabo-ration (A)	3 - Appreciate the perspectives of others 6 - Handle interpersonal situations constructively		2 - Discern and understand emotion of other 4 - Empathic involvement in other
Task Per-formance (C)	5 - Make responsible decisions 2 - Set and achieve positive goals	5 - Responsible decision-making	
Emotion Regulation (N)	1b - Manage emotions	2 - Self-management	6 - Coping with aversive or distressing emotions 8 - Accept own emotions (emotional self-efficacy)
Open-mindedness (O)	1a - Recognize emotions	1 - Self-awareness	1 - Aware of own emotional state 3 - Use local emotion vocabulary correctly 5 - Appreciate emotion experience differs from expression

Note. If the model in this column does not include a relevant construct, the space is left empty. Labels given in the first column are the terms in the socio-emotional framework proposed here. In parentheses, we provide the acronym abbreviations of the old labels for the Big Five (E = Extraversion; A = Agreeableness; C = Conscientiousness; N = Negative emotionality; O = Openness).

Researchers have commented on this lack of standard constructs and instruments. For example, Furlong et al. (2007) commented that “Assessments have begun to be developed, but have not had time to fully mature” (p. 2). When they examined the structure of the Behavioral and Emotional Rating Scale (BERS-2; Epstein, 2004), which is now more commonly used to evaluate the outcome effectiveness of intervention studies, they noted, “The BERS was developed using a mixture of professional judgement and empirical procedures; however—it had no prespecified theoretical foundation or set of psychological constructs” (p. 2). The *Intrapersonal Strength* subscale measures a youth’s outlook on his or her competence and accomplishments (e.g., “I believe in myself”). The *Affective Strength* subscale measures the ability of a child to accept affection from others and express feelings towards others (e.g., “It’s okay when people hug me”). The *Interpersonal Strength* subscale measures a youth’s ability to control his or her emotions or behaviours in social situations (e.g., “I can express my anger in the right way”). The *School Functioning* subscale measures competence in school and classroom tasks (e.g., “I complete tasks when asked”). The *Family Involvement* subscale measures a child’s participation in and involvement with his or her family (e.g., “My family makes me feel wanted”).

On the basis of their extensive review of social and emotional learning programs, Durlak et al. (2011) concluded that socio-emotional intervention researchers tend to focus on five broad competency sets, including self-awareness, self-management, social awareness, relationship skills, and responsible decision-making. As shown in Table 4.4, these attributes, as well as the six defined by Elias et al. (2007), map reasonably well onto the socio-emotional Big Five shown in the first column of Table 4.4. For example, Elias et al.’s domain 3 (Appreciating the perspective of others) should be well represented by the facets of our Collaboration (Agreeableness) domain, especially by our facet 2 “Respect for others, empathy, tolerance, fairness.” Similarly, Durlak et al.’s first domain is Self-awareness (which involves the ability to introspect and take note of one’s inner experiences, like thoughts and feelings) should be well-represented by our Interest & awareness (Openness) domain, especially by facet 4 “Self-reflection, self-awareness, consciousness, self-actualization, authenticity.”

The Emotional Competence Approach and Its Major Measures

The *emotional competence* approach originated in developmental and clinical analyses of what a child needs to *learn to become an emotionally and socially competent adult*. One major theoretician and proponent of the competence approach is Saarni (e.g., 1999; 2011) and her conceptual work has been quite influential. She analysed emotional functioning from the perspective of how well it serves the adaptive and instrumental goals of the individual and then defined emotional competence as a set of affect-oriented behavioural, cognitive and regulatory skills. Simply put, the child needs to learn both what it means to *feel* something and to *do* something about those feelings.

As shown in Table 4.4, Saarni (e.g., 1999; 2011) postulated 8 basic skills she considered prerequisites for emotional competence: (1) Awareness of one’s own emotional state; (2) Skills in discerning and understanding the emotions of others; (3) Skill in using the common vocabulary of emotion and expression; (4) Capacity for empathic and sympathetic involvement in others’ emotional experiences; (5) Skill in realizing that inner emotional states need not correspond to outer expression; (6) Capacity for adaptive coping with aversive or distressing emotions by using self-regulatory strategies that ameliorate the intensity or temporal duration of such emotional states; (7) Awareness that relationships are defined by emotional genuineness of expressive display and reciprocity; and (8) Capacity for emotional self-efficacy (i.e., individuals can accept their own emotional experience and view themselves as generally feeling the way they want to feel). Although not necessarily accepted by all researchers in the field, these 8 competencies provide a formidable set of complex socio-emotional skills for researchers to define and measure.

Note that competency (6), *capacity for adaptive coping* with negative emotions like sadness, anxiety, or anger, would be most directly relevant for individual differences in *emotional outcomes* in the child. The other competencies specify important processes that may help (rather than hinder) the enactment of effective regulatory strategies in both emotional and social situations. For example, if an individual has no competency (1), *awareness of his or her current emotional state*, then the individual would hardly proceed with (6) attempts to cope with those emotions. Similarly, lack of emotional self-awareness is likely to interfere with (7) genuine emotional expressions and reciprocity in social contexts.

Self-report Questionnaires of Emotional Competencies

Although the emotional competence approach has not generated a commonly accepted taxonomy and measurement model, it has generated several individual difference measures, mostly for adults. These measures and their strengths and weakness have been reviewed twice in the last decade (see John and Gross, 2007; John and Eng, 2014), and we therefore provide an abbreviated version here. For more details, tables, and references, we refer the reader to these original sources.

The Generalized Expectancy for Negative Mood Regulation Scale (NMR), developed by Catanzaro and Mearns (1990), was one of the earliest measures. It focused on individuals' *beliefs* that some "behaviour or cognition will alleviate a negative state or induce a positive one" (p. 547), and asks participants to indicate the extent to which they believe that their attempts to alter their negative moods will work. Many of the items focus on ways to avoid negative emotions. Thus, the measure has been criticized for equating mood regulation with the avoidance of negative affect (e.g., Gratz and Roemer, 2004); simply avoiding negative emotion is assumed to be an indication of effective regulation, as shown by items such as "When I'm upset, I believe that I can forget about what's upsetting me pretty easily" versus "When I'm upset, I believe that I won't be able to put it out of my mind" (reverse scored).

Consistent with the emotional competence perspective, Salovey, Mayer, Goldman, Turvey, and Palfai (1995) aimed to understand the reflective or "meta" processes that accompany many mood states. These "meta-mood" processes capture how individuals reflect on their feelings, including how they monitor, evaluate, and regulate them (Mayer and Gaschke, 1988). Salovey et al. (1995) assumed that emotions serve as an important source of information for the individual, and that individuals differ in how skilled they are at processing this kind of information, particularly in "their understanding of and ability to articulate their affective states" and in their ability to "regulate such feelings and use them adaptively to motivate behavior" (p. 147). The *Trait Meta-Mood Scales (TMMS)* were designed to measure stable and general attitudes about moods and the degree to which individuals attempt to manage (or repair) mood experiences. The TMMS measures Saarni's construct *awareness of one's own emotional state* in terms of two scales: (a) the tendency to attend to one's moods and emotions (*attention*) and (b) to discriminate clearly among them (*clarity*). The third scale aims to assess efforts to *repair* one's emotional state if needed. These individual differences were considered "fundamental to the self-regulatory domain of emotional intelligence" (Salovey et al., 1995, p. 147).

The TMMS Attention scale refers to paying close attention to feelings, accepting feelings, valuing them positively, and letting oneself experience them fully and intensively, using items such as "I often think about my feelings" versus "I don't think it's worth paying attention to your emotions or moods" (reversed-scored). As expected, the Attention scale correlated with the Private Self-consciousness Scale, which measures awareness and attention to private aspects of the self (like thoughts and feelings), which is represented in the Big Five personality taxonomy by the Openness to feelings facet; this facet is also well-represented in Table 4.3. The TMMS Clarity scale assesses feeling at ease and clear about one's emotions, as contrasted with a deep and troubling confusion about one's emotions and what they mean. This should

interfere with effective mood repair and, indeed, low clarity was related to vulnerability to negative affect, distress, and depression. The TMMS Repair scale assesses attempts to improve negative mood by thinking positively and taking an optimistic (rather than pessimistic) attitude more generally. Item examples include “Although I am sometimes sad, I have a mostly optimistic outlook” and “I try to think good thoughts no matter how badly I feel,” as contrasted with “Although I am sometimes happy, I have a mostly pessimistic outlook” (reverse-scored). This scale seems conceptually similar to coping measures and was found to correlate substantially with low vulnerability to distress and depression as well as with greater optimism. More generally, these findings are consistent with the structural model proposed in Table 4.4: competencies or skills in emotion regulation, like mood repair, should lead to better emotional outcomes when individuals face aversive or stressful situations. Again, these socio-emotional characteristics are well-represented in the model laid out in Table 4.3.

Gratz and Roemer (2004) followed Saarni’s (1999) approach to emotional competencies, and were influenced by Salovey et al.’s TMMS scales, as well as by the older NMR (p. 44). However, they deviated from the “strengths” or “skill” based approach and instead devised a measure they called *Difficulties with Emotion Regulating Scale (DERS)* during times of distress. Specifically, their items used the format of the older NMR, and all items begin with the sentence stem “When I’m upset, I... .” This focus on global negative affect (upset) is a feature that both the NMR and the DERS share with measures of coping. Even though the DERS has six subscales, they are substantially inter-correlated and are often aggregated into a single overall dysregulation score, which correlates substantially with various indicators of negative affect, psychopathology, and low well-being (e.g., Weinberg and Klonsky, 2009).

Example of a Situational Judgment Test: Emotional Intelligence Test (MSCEIT)

It should be clear by now that the emotional competence approach conceptualizes emotion regulation in terms of a number of specific *abilities*. Yet, all the measures we have discussed so far have used self-report questionnaire methodology, asking about self-perceptions (including self-efficacy beliefs) and about typical experiences and behaviours. In fact, we have seen that some of these “competence” scales seem indistinguishable from coping styles. This is hardly a compelling way to assess constructs defined as abilities such as intelligence, which psychologists measure with maximum performance tests of the behaviour or process in question.

Mayer, Salovey, and Caruso (2002) acknowledged this methodological inconsistency and undertook the challenging task of constructing an “emotional intelligence test” (abbreviated MSCEIT) scored objectively in terms of correct and incorrect answers. They define emotional intelligence as a set of skills involved in the processing of emotion-relevant information. Here we address, as an example, only emotion management ability, which is the most relevant of the MSCEIT components and is defined as the capacity to reduce, increase, or maintain particular emotions in both oneself and other people. The tasks used to measure these abilities follow the format known as a *situational judgment test*; these tests aim to assess the ability to choose the most appropriate action from a pre-specified set of options and are typically used with adults in workplace and job selection contexts (but see a recent application to collaboration skills in adolescents by Richard Roberts and colleagues, 2012). The MSCEIT requires respondents to react to hypothetical scenarios and evaluate the effectiveness of various behaviours and subjective construals for emotion management purposes. For example, participants are asked to judge the effectiveness of strategies to help a friend enhance a joyful mood or reduce feelings of sadness.

The test itself is owned by a commercial “test publisher (who) does not authorize reproduction of actual test items” (e.g., Lopes et al., 2005, p. 114). This has been an impediment to research on the MSCEIT and its scientific evaluation has been hindered; thus, , we reprint below the only two abridged item examples available, both from Lopes et al. (2005, pp.114-115). Each item consists of a vignette paired with separate response options:

“Debbie just came back from vacation. She was feeling peaceful and content. How well would each action preserve her mood? (1) She started to make a list of things at home that she needed to do. (2) She began thinking about where and when to go on her next vacation. (3) She called a friend to tell her about the vacation . . .”

“Ken and Andy have been good friends for over 10 years. Recently, however, Andy was promoted and became Ken’s manager. Ken felt that the new promotion had changed Andy in that Andy had become very bossy to him. How effective would Ken be in maintaining a good relationship, if he chose to respond in each of the following ways? (1) Ken tried to understand Andy’s new role and tried to adjust to the changes in their interactions. (2) Ken approached Andy and confronted him regarding the change in his behavior.”

These examples are from the fourth (Managing Emotions) “branch” of the MSCEIT, which consists of two distinct tasks. Five vignettes measure *ability in emotion management*, and each describes a person (like Debbie above) who is experiencing a mood or emotion. For each of the 5 vignettes, the respondent rates (on a 5-point scale) how effective four different actions would be for obtaining a specified effect on the person’s experience (here, to preserve Debbie’s good mood), yielding a total of 20 separate ratings. The second task measures *emotional relationship abilities* and consists of three vignettes describing relationships between persons (like Ken and Andy above). In each vignette, the respondent rates how effective three different actions would be *to maintain a good relationship* between the persons, for a total of 9 separate ratings. Each of the 29 individual ratings is scored according to a normative effectiveness rating provided by a panel of emotion experts or the group consensus.

Although abbreviated, these two examples are very instructive. First, the total emotion management ability score includes more than 30% of the ratings that do not involve emotional but relational skills, raising questions about content validity. Second, each vignette and action includes a lot of detailed contextual information specific to that rating, which adds error and lowers inter-item correlations and thus reliability; with 29 ratings aggregated into the total score, reliability in this study was a modest .63, and that is higher than in other studies (see Føllesdal and Hagtvat, 2009, for a thoughtful psychometric analysis and critique). Third, as the MSCEIT authors readily acknowledge, these vignette ratings do not actually measure individual differences in skilful or effective regulation scored or observed objectively in an emotional situation; instead they tap the individual’s *knowledge, and capacity to reason*, about emotions and emotional situations (e.g., Lopes et al., 2005, p. 114). Fourth, the emphasis on knowledge and complex reasoning processes is likely to introduce correlations with measures of other abilities, creating discriminant validity problems. Fifth, there are interpersonal themes even in the emotion management vignettes (e.g., calling a friend to share one’s mood nature and thus capitalize on the experience) and thus performance on these items may yield surprising correlations with personality variables, again introducing potential problems with discriminant validity. In response to these discriminant validity concerns, the test authors and their collaborators have argued that the MSCEIT scores predict social, emotional, and leadership outcomes even when intelligence and broad personality traits are controlled. So far, however, many researchers have remained unconvinced; the MSCEIT, and emotional intelligence research more generally, is viewed with scepticism among researchers (e.g., Landy, 2005; Joseph and Newman, 2010).

For example, Lopes et al. (2005) obtained 8 criterion measures (e.g., interpersonal sensitivity; socio-emotional competence; friendship nominations) with self-ratings or peer nominations. When the Big Five Inventory personality scales (John et al., 2008) were controlled, the MSCEIT emotion-management ability scale still significantly predicted two of these criteria. However, with 2 out of 8 correlations significant, it is hard to say whether the predictive-validity goblet is a quarter full or three quarters empty. Of greater interest to personality researchers is a finding not highlighted by the authors: by far the highest correlation was not found in predicting any of the 8 socio-emotional outcome measures but the Big Five dimension of Agreeableness; the *r* of .40 is a very substantial correlation once compared with the modest reliability of

.63 of the MSCEIT scale in this study. Again, we have defined several facets relevant to reciprocal collaboration that should well capture any individual differences related to Agreeableness in the MSCEIT.

In conclusion, the MSCEIT, though an admirable and conceptually interesting undertaking, has not proven the decisive fix for the self-report measures of emotional “competencies” that have come before it. Even though as outsiders we do not know much about the inner workings of the MSCEIT, it seems unlikely that scores on its 5 emotion management vignettes and the 3 relationship vignettes can provide the conceptual building blocks needed to construct a comprehensive measure of socio-emotional abilities.

More generally, the conceptual richness, reach, and resulting complexity of the emotional competence approach (e.g., Saarni, 1999) may be a strength as well as a major limitation. It may include too many cognitive, behavioral, self-perception, and emotion perception processes under one broad rubric. Fewer constructs, more narrowly delineated distinctions, and tighter links between construct definitions and actual measures may prove a fruitful avenue for future research. Nonetheless, we were able to draw on this approach for a better understanding of our social-emotional constructs in Table 4.3 and to supplement the conceptual framework for the OECD longitudinal study.

We now turn to two other concepts that educational psychologists have developed over the past decades and that are often mentioned in the context of social-emotional learning and 21st century skills, namely metacognitive skills and learning styles. Conceptually speaking, however, these two are probably better described as social-cognitive instead of social-emotional skills. Similarly, an important social-cognitive belief construct that we plan to add to the model in Table 4.3 is Carol Dweck’s implicit theories concept (e.g., Dweck et al., 1993); she contrasts entity (fixed and unchanging) beliefs about ability (and intelligence) with incremental (growth-oriented) beliefs that abilities can change and grow; growth beliefs have been shown to predict much better academic and well-being outcomes than beliefs that abilities are fixed and unchanging.

Meta-cognitive skills

Schraw and Dennison (1994, p. 460) define meta-cognition as: “The ability to reflect upon, understand, and control one’s learning”, or differently phrased ‘thinking about your thinking’ (Flavell, 1979) in the context of learning. The metacognition concept taps into higher-order mental processes referring to knowing what strategies work best for learning and how and when to activate these strategies (metacognitive knowledge), but also to the capacity to regulate these skills reflected in activities such as planning, information management strategies, comprehension monitoring, debugging strategies, and evaluation (subsumed under metacognitive regulation). Schraw and Dennison (1994) found that the knowledge and the regulation components are distinct components though are associated in the .40 to .50 range. There is support for the predictive validity of metacognition for academic performance (.21; Coutinho, 2007), and there is evidence that it mediates the relationship between mastery goal setting and academic success (Coutinho, 2007).

Schraw and Dennison (1994) developed the Metacognitive Awareness Inventory (MAI), a 52-item self-report instrument to assess metacognitive awareness and its components. The MAI has a two-factor structure, including of facets. A sample item is written between parentheses, to be in a better position to examine coverage by the proposed social-emotional framework. The first ‘knowledge of cognition’ dimension includes three facets: (a) Declarative knowledge, described as “knowledge about one’s skills, intellectual resources, and abilities as a learner” (p. 474) (“I am good at organizing information”), (b) Procedural knowledge, i.e. knowing how to apply learning strategies (“I try to use strategies that have worked in the past”), and (c) Conditional knowledge, i.e. knowing when and why to use particular learning skills (“I use different learning strategies depending on the situation”). The second factor, ‘regulation of cognition’ has five facets: (a) Planning, i.e. the process of planning and preparation,

goal-setting and allocation of investments (“I set specific goals before I begin a task”), (b) Information Management, i.e. the efficient and timely processing of available information (“I slow down when I encounter important information”), (c) Monitoring, i.e. controlling and assessing the learning process and use of strategies (“I ask myself periodically if I am meeting my goals”), (d) Debugging, i.e. correcting and redirecting learning strategies when necessary (“I ask others for help when I don’t understand something”), and (e) Evaluation, i.e. analysing and reflecting on the learning performance and the obtained result after learning (“I summarize what I’ve learned after I finish”). A content analysis of the 52 MAI-items showed that the majority of items could be described as indicators of Conscientiousness and Openness to experience, though contextualized in a learning situation.

The construct of metacognition is represented in a number of 21st century frameworks, including other frameworks discussed below, such as those proposed by the Center of Curriculum Redesign (CCR) and by DeSeCo (Definition and Selection of Key Competencies), in which reflection is at the core concept.

Learning styles

Students show a huge variety in how they perceive and learn, and this variability has been represented in models and theories on learning styles or learning approaches. Learning style is usually narrowly conceived, i.e. as a combination of different learning activities or as a learning strategy, but when depth of information is emphasized, the concept ‘approach’ is mostly used. Different learning style instruments are around which are fairly similar to one another (Furnham, 1996). Honey and Mumford’s (1982) Learning Style Questionnaire (LSQ) was based on Kolb’s (1976, 1984) theory, and distinguished four different types of learning: activists, reflectors, theorists, and pragmatists. Activists are open-minded and get fully and enthusiastically engaged in new experiences, whereas ‘reflectors’ like to stand back and (re-)evaluate the different elements before deciding or acting. Theorists adapt, analyse and integrate distinct facts into coherent theories, but ‘pragmatists’ check out whether something new works in practice. Another measure is the Approaches to Studying Inventory (ASI) developed by Entwistle and Tait (1995) distinguishing among level of engagement and depth of processing when learning. This inventory assesses deep (intention to understand, relating ideas, use of evidence and active learning), surface (intention to reproduce, unrelated memorizing, passive learning and fear of failure), and strategic (study organization, time management, alertness to assessment demands and intention to excel) learning approaches.

The association between learning strategies and the Big Five dimensions were first described by Diseth (2003a) among samples of psychology and philosophy students. Across the samples, the deep learning approach was associated with Openness (.46 and .54), the surface approach positively with Neuroticism (.42 and .49) and negatively with Openness (-.25 and -.21), and the strategic learning approach with Conscientiousness (.55 and .62). Despite this conceptual and empirical overlap, learning styles added to the prediction of educational attainment. Komarraju, Karau, Schmeck and Avdic (2011) found that learning styles explained an additional 3% of GPA variance on top of the 14% already explained by the FFM (see also Rosander and Bäckström, 2012). More important, several studies have shown that specific learning styles mediate the relationship between the Big Five and examination grades (Komarraju et al., 2011, for Openness, and Diseth, 2013b, for Openness, Neuroticism and Conscientiousness), suggesting they are better conceived as characteristic adaptations resulting from more basic cognitive and non-cognitive tendencies.

4.3. Malleability

A number of observational and intervention studies provide evidence on malleability of the proposed constructs.

Observational studies

Research demonstrating that personality traits show substantial plasticity and continue to develop in adulthood is now widely available (e.g., Helson, Kwan, John, and Jones, 2002; Srivastava et al., 2003; for a review, see Roberts, Walton, and Viechtbauer, 2006). Extensive observational studies have demonstrated that throughout early and middle adulthood, many people increase in what has been called *psychosocial maturity*, that is, they increase in conscientiousness and agreeableness, and they decrease in negative affect. However, much of that research has been conducted only at the broad level of the Big Five domains; recent research including facets (Soto and John, 2014) suggests that one developmental pattern does not always hold for all the facets in a domain. For example, among the conscientiousness facets, self-discipline increased substantially from age 20 to 60 whereas orderliness did not increase much at all.

Observational research on naturally occurring personality development in children and adolescents is just beginning to hit its stride (see the recent special issue of the *European Journal of Personality*, edited by Denissen, 2014). This slower start occurred, in part because research on children faces even greater hurdles than longitudinal studies of personality change in adults (see the final section of this report for a discussion of some of those issues). With younger children, self-reports cannot be used as an efficient method to collect data; assessment instruments have to be made age-specific, so that they are appropriate for the age-specific emotional and behavioural repertoire of the child and then the adolescent, making it more difficult to compare developmentally as instruments change; and changes occur more rapidly than in adults, necessitating yearly assessments whereas in adult life-span studies assessments may be limited to every 5 or even 10 years (George, Helson, and John, 2011). Again, most of the available studies of children have focused on only the Big Five domains (see Roberts et al., 2006, for a review), thus cannot speak to the more differentiated facet structure of the framework considered here.

One recent study has used the same instrument (the BFI) from late childhood (age 10) to late adolescence (age 20), thus making mean-level comparisons across these age groups easier; age differences were examined for both Big Five domains and facets, in a large sample recruited via the internet (Soto, John, Gosling, and Potter, 2011). Clear evidence emerged for substantial age differences throughout this difficult-to-measure development period, with a curvilinear pattern. After age 11, the data showed that on average, socio-emotional functioning was challenged by the onset of adolescence; however, by age 15, both girls and boys had begun to recover from the “Sturm und Drang” of early adolescence and showed positive age trends on the facets related to psychosocial maturity (higher agreeableness and conscientiousness) all the way to age 20.

Gender differences also developed during this adolescent period. At age 10, boys and girls did not differ in negative affect but then girls quickly increased to the elevated levels typical of young adult women (by age 15). In contrast, boys stayed stable overall and increased somewhat in self-confidence (i.e., lower anxiety levels) all the way to age 20. On the other hand, gender differences in agreeableness were already apparent in late childhood, with girls scoring higher than boys in each age studied. One needs to be careful before generalizing from a single (if large) sample study in one country. However, similar effects are being found in a cross-cultural sample assessed starting from age 12, with the gender difference in negative affect emerging again only in adolescence, and replications are under way, for example, in Brazil.

The critical question for future research, however, involves individual differences in change trajectories, which can only be studied in a longitudinal design. Specifically, we expect that not all

adolescents will show the age-typical deterioration in socio-emotional functioning; worse, not all who did show that deterioration will recover from it as quickly as the normative data suggest, as problems such as juvenile delinquency emerge during this time and can potentially create a longer negative developmental dynamic.

The OECD Study on Skill Development will be longitudinal, and should include at least yearly measurement points to be able to capture such non-linear patterns of rapid change during the school years. The study will be designed to learn about social-emotional skill development trajectories and how these relate to a broad range of outcomes, some of which emerge during that age period. At the same time, developmental contexts during these trajectories will be prospectively studied, with the individual countries in which the studies are conducted serving as macro-environmental factors. The primary questions of this investigation will hence be centred on within-country comparisons across time.

Intervention Studies

One of the reasons why the Longitudinal Study of Social and Emotional skills in Cities is needed is that, at this point, there is no compelling research available where the development of social-emotional skills is studied longitudinally in a natural context, across a substantial time-interval, using adequate measures for which measurement equivalence has been demonstrated, and relying on multiple informants. The reason for the paucity of systematic research is, as we noted earlier, the lack of consensus on how to define social-emotional skills and construct a consensual taxonomy representing their features and content.

In contrast, there are a multitude of studies examining the impact of different kinds of school-based interventions to enhance students' social and emotional learning. These programs usually aim to either increase particular socio-emotional skills (e.g., peaceful conflict resolution) or influence a specific subset of the outcomes to be targeted in the Longitudinal Study of Skill Development in Cities, including positive social behaviours, conduct problems, emotional distress, psychological well-being, physical health, and academic performance.

A number of impressive meta-analyses have been conducted (Durlak, Weissberg, Dymnicki, Taylor, and Schellinger, 2011; Park-Higgerson, Perumean-Chaney, Bartolucci, Grimley, and Singh, 2008; Sklad, Diekstra, De Ritter, Ben, and Gravesteyn, 2012) examining the impact of such interventions, with special attention for important moderators. Durlak and colleagues conducted an impressive meta-analysis on the impact of school-based social-emotional learning programs published before 2007, summarizing findings obtained from kindergarten to high school and reporting on a total sample of $N=270,034$. As moderators they included whether programs were (a) run by expert/consultants or by teachers themselves, (b) organized at the level of the classroom only versus at the classroom and the school level, (c) developed according to SAFE (sequenced, active, focused and explicit) or non-SAFE criteria, and finally whether (d) implementation problems for the programs were reported or not. Overall, small to moderate intervention effects were reported for attitudes, positive social behaviour, conduct problems, emotional distress, and academic performance, with moderate effects reported for the malleability of social-emotional skills. Across these criteria, programs implemented by teachers showed significant impact, suggesting that teacher-based programs are effective, and that one does not necessarily need external personnel/consultants in the classroom to achieve results. Only programs designed according to SAFE criteria, this means an active involvement of pupils, following a sequenced, focused and explicit program, demonstrated effectiveness. Finally, only interventions for which no implementation problems were reported or not mentioned turned out to be effective. A constraint of this meta-analysis was that 53% of the source of the outcome data were child-reported.

A second major meta-analysis on the subject has been conducted by Sklad and collaborators (2012), reporting effects of 75 universal school-based intervention programs for which the data were

published between 1995 and 2008 with an average reported intervention sample size of $N=543$ (range 13 to 8280). Skald et al.'s meta-analysis provides an excellent follow-up on Durlak et al.'s review, because they also included 16 non-American based studies (21% of the total meta-analysis) and investigated immediate and delayed outcomes. The majority of the reviewed studies had a post-test between 0 to 6 months (73.3%), 36% of the studies had a follow-up between 7 and 18 months and for 21.3% follow-up data were available after 19+ months. Again here, the outcome measurement relied chiefly on self-reports (60% of the programs) and for 73.3% of the programs, no intervention manual was available, making it difficult to really study the content of interventions.

Skald and colleagues (2012; Table 4.4) found substantial evidence indicating an improvement but the effect sizes varied by domain targeted for intervention; d effect size estimates for immediate effects were .70 for socio-emotional skills, .46 for positive self-image, .46 for immediate academic achievement, -.43 for antisocial behaviour, .39 for prosocial behaviour, -.19 for mental disorders, and -.09 for substance abuse. In other words, socio-emotional skills were most malleable in intervention contexts, whereas mental disorders and substance abuse were least affected by the interventions. As one would expect, effect sizes at a later follow-up decreased substantially for all outcomes, with effect sizes reduced to .26 for academic achievement, -.20 for antisocial behaviour, -.10 for mental disorders, .07 for positive self-image, .12 for prosocial behaviour, .07 for social-emotional skills and -.18 for substance abuse. The authors concluded from these data that, despite large immediate gains, long-term effects were small, with the average program participant still outperforming the average non-participant by 5%.

Additional key findings were that programs with a duration of less than a year had more impact on social skills than those that had a longer time-frame; also a smaller number of sessions (less than 20 sessions) turned out to be more effective. Intervention impact on social skills was equally large in primary and secondary school, whereas effectiveness to reduce antisocial behaviour was strongest in primary school. These findings suggest that antisocial behaviours are better tackled early on at school, whereas there is equal room for improvement of social skills across both primary and secondary school. Teachers in Sklad's analysis further turned out to be as effective as non-teachers to run programs, confirming Durlak's (2011) conclusion that teachers can successfully implement these programs. Finally, interventions' impact on social skills seems to be equal in North-American samples versus studies conducted outside of North America, suggesting that malleability generalizes across societies.

Besides these meta-analyses targeting a broad range of outcomes, there is also a wide range of studies, including randomized control trials, on reducing aggressive behaviours (e.g. Park-Higgerson, et al., 2008), and focusing on antisocial personality (Scott, Briskman, and O'Connor, 2014), oppositional defiant disorder (ODD) (Scott, et al., 2014), and conduct disorder (CD). There are effect evaluations examining broad and intensive clinical programs, often also working with parents (Scott, et al., 2014), broad and intensive community versus clinical programs (Kolko et al., 2009), short (reduced) programs) and the effect of organizing booster sessions (Lochman et al., 2014) to maintain long-term effects of interventions.

5. COHERENCE WITH OTHER FRAMEWORKS AND EDUCATIONAL GOALS

In this final review section, we examine how the framework proposed here maps onto the three major domains of functioning specified by the OECD (Managing emotions; Working with others; Achieving goals). In addition, we show links between the Big Five derived measurement framework and other approaches that aim to define the major *goals or objectives of education*, including the approach proposed by the *Collaborative for Academic, Social, and Emotional Learning* (CASEL), which is closely related to the socio-emotional learning approaches by Durlak, Elias, and colleagues reviewed above, the integrative proposal for educational goals by the Center for Curriculum Redesign (CCR), and finally the framework adopted by the *Knowledge is Power Program* (KIPP) schools.

Table 5.1 provides an integration of different frameworks that have been suggested previously for the social-emotional skill domain. The table is not meant to be exhaustive, and primarily focuses on frameworks that have been reviewed earlier in this report. It starts from the common framework offered for the OECD's Longitudinal Study of Skill Development in Cities (OECD, 2015), proposing the skill domains of 'Managing emotions', 'Working with others', and 'Achieving goals'. Such broad grouping of social-emotional skills is particularly useful for conceptual and communicative purposes to link skills with potential outcomes, though these domains are too general to be applicable at an operational level to track development from childhood to young adulthood. The Longitudinal Study of Skill Development in Cities exactly needs a more fine-grained proposal representing the commonality across constructs suggested in different social-emotional skill frameworks, but also needs to cover those skills and constructs that are necessary to understand developmental trajectories and explain outcomes. In addition, the suggested system should be sensitive to cultural differences and also embrace constructs that may help to explain trajectory and outcome variation across cultures.

As described earlier, the currently best-researched taxonomy of individual differences that comes closest to the domain of social-emotional skills is the Five-Factor taxonomy of personality. Personality psychologists now agree that five main dimensions represent the core qualities underlying personality differences. Although consensus at the lower-order facet level in the Five-Factor Model hierarchy has not yet been established, Table 4.3 shows that there is starting convergence also at that level. At the same time, these five dimensions also describe how people interact with and adapt to the environment they live in. It should hence not come as a surprise that these five dimensions also have a cardinal position in an analysis of social-emotional or 21st century skills. Emotional stability directly taps into the domain of 'Managing one's emotions', Extraversion and Agreeableness describe how we 'get along, engage, and work with others', whereas Conscientiousness and Openness to experiences are respectively about 'getting things done/achieving goals', and 'being explorative and innovative'. These different adaptive functionalities map well onto the OECD's framework (column 1).

Table 5.1. Mapping Different Frameworks Proposed for Socio-Emotional Skills and for Goals of Education

OECD	Skill equivalents here	Proposed constructs here	CASEL	CCR	KIPP Schools
Managing emotions	Emotion awareness skills Emotion acceptance skills	Stress resistance Self-esteem Emotional control	Self-management	Mindfulness (Q)	Self-control
	Emotion reappraisal skills Emotion modification skills	Self-compassion Self-confidence Fear-of-happiness			
Working with others	Assertiveness skills Presentation skills Social contact skills Leadership skills	Assertiveness Enthusiasm Social approach & connection		Courage (Q)	Zest/optimism
	Collaboration skills	Interdependent self-construal Compassion Trust	Relationship skills	Leadership (Q)	Collaboration (S)
	Communication skills	Relationship harmony Respect for others	Social awareness	Communication (S) Ethics (Q)	Gratitude Social intelligence
Achieving goals	Responsible decision-making skills Goal setting skills Task engagement skills	Responsibility Goal orientation Task initiation Self-discipline Organization	Responsible decision-making	Resilience (Q)	Grit
	Creativity skills Appreciation skills Self-reflective skills Critical thinking skills	Creative Imagination Intellectual Curiosity Aesthetic interests Self-reflection/awareness Autonomy/Independence	Self-awareness	Creativity (S) Curiosity (Q)	Curiosity
				Critical thinking (S)	

5.1. Extensions to Other Frameworks

To further illustrate the FFM's dimensions' status as constructs referring to the core qualities represented in the amalgam of skills, learning objectives and attitudes, that are proposed as standards of aspiration in the 21st century, we classified the constructs proposed by two key players in the social-emotional learning field in Table 5.1, in the columns 4 and 5.

The influential Collaborative for Academic, Social, and Emotional Learning (CASEL; www.casel.org) from the University of Illinois at Chicago has suggested five competency areas that should be advocated in social-emotional learning programs, i.e. Self-awareness, Social-Awareness, Self-management, Relationship skills, and Responsible decision-making. Three of these competency areas are content-wise consistent and align unequivocally with specific FFM dimensions. Social awareness (understanding and empathy, taking others' perspectives) and Relationship skills (working in teams, positive relationships, reciprocity, conflict handling) are clearly related to the Agreeableness domain, describing individual differences in the quality of social interactions with persons. Self-awareness (recognizing one's emotions, self-confidence, accepting limits, but also recognizing strengths) is without doubt related to the FFM Emotional Stability dimension.

Self-management (being able to control and regulate emotions to achieve goals, conscientiousness and perseverance) is a more hybrid competency domain from an FFM perspective, primarily related to Emotional Stability, but also to Conscientiousness. Finally, Responsible decision-making (assessing risks and making deliberate decisions, respecting others) is more at the intersection of Conscientiousness and Agreeableness, so we have classified this skill domain in the Conscientiousness area, but also closely to the Agreeableness dimension. The classification of the CASEL framework within Table 5.1, illustrates its strong background in positive psychology, whereas it seems to lack competency areas that have to do with extraverted qualities and openness to experience, two other key dimensions to engage with the world outside.

The integrative set of skills and qualities proposed by the Center for Curriculum Redesign (CCR, 2015) are described in column 5. These concepts can be classified within the FFM framework when considered at the level of their overarching labels. Mindfulness (e.g., self-awareness, tranquillity) aligns with Emotional stability, whereas Courage (e.g., bravery, energy) and Leadership (e.g., charisma, assertiveness, responsibility) refer to "getting ahead" and are thus conceptually related to Extraversion. Ethics (e.g., benevolence, compassion, honesty) associates with Agreeableness, whereas Resilience is defined by CCR in terms of perseverance, grit, and self-discipline, and thus maps primarily onto Conscientiousness. Curiosity importantly taps into the Openness to experience domain. Where do the "Four Cs" in the CCR concepts, namely Collaboration, Communication, Creativity, and Critical thinking fall? Collaboration and communication belong in the Agreeableness domain, whereas creativity and critical thinking belong in the Openness domain. Overall, it appears that when evaluated from the perspective of our OECD measurement framework, the CCR concepts are more comprehensive than, for example, the CASEL framework because they provide a better coverage of the Openness domain.

The final framework shown in Table 5.1 is the Knowledge is Power Program (KIPP). In association with Dr. Duckworth (www.kipp.org/our-approach/character), they proposed seven constructs in their character strength approach, including zest, grit, optimism, self-control, gratitude, social intelligence and curiosity. As can be seen in the last column of Table 5.1, their constructs can be conceptually classified across the five dimensions defined by the Big Five.

By extension, similar conceptual classifications can be made for other frameworks that have been developed by scholars and interest groups in the social-emotional learning area. The Partnership for 21st Century Skills (www.p21.org/our-work/p21-framework) distinguished among Life and Career Skills,

Learning and Innovation Skills, and Information, Media, and Technology Skills. Life and Career Skills include “flexibility and adaptability”, “initiative and self-direction”, “social and cross-cultural skills”, “productivity and accountability”, and “leadership and responsibility”. The Learning and Innovation Skills group encompasses “creativity and innovation”, “critical thinking and problem solving”, and “communication and collaboration”. Again, this skill set sorts across four of the five FFM categories, except for the Emotional Stability domain. P21 further distinguishes Information, Media and Technology Skills, including ‘information literacy’, ‘media literacy’, and ‘ICT literacy’. Although part of the skills in this group (e.g., flexibility and critical thinking) may be related to openness, they also tap into more functional skill domains, related to specific technologies.

In association with the OECD, the Swiss-led DeSeCo Project (Definition and Selection of Key Competencies) aimed to build a competency framework for measuring the competence level of young people and adults across different countries. The DeSeCo Model groups its key competencies into three broad categories: Using tools interactively, Interact in heterogeneous groups, and Act autonomously. In addition, they underscore the importance of reflectiveness, defined as the ability to deal with change, learn from experience and critical thinking. Using tools interactively is further split into: ‘use language, symbols and texts interactively’, ‘use knowledge and information interactively’, and ‘use technology interactively’; Interacting with heterogeneous groups includes ‘relate well to others’, ‘co-operate and work in teams’, and ‘manage and resolve conflicts’; Acting autonomously refers to ‘act within the big picture’, ‘form and conduct life plans and personal projects’, and ‘defend and assert rights, interests, limits and needs’.

Finally, the Strive Partnership (www.strivetogether.org) distinguishes among the social-emotional competencies of ‘Academic self-efficacy’, ‘Growth mindset and mastery orientation’, ‘Grit or perseverance’, ‘Emotional competence’, and ‘Self-regulated learning and study skills’. The factors specified by these theories and others in the literature, such as Self-Determination Theory (Deci and Ryan, 2002), Goleman’s Emotional Intelligence (1995), or Blair’s (2002) work on social and emotional competencies are represented by the facet constructs, either singly or in combination, proposed in this framework.

5.2. Cultural Sensitivity

As discussed earlier, Table 5.1 also includes several constructs hypothesized to be important to understand cultural differences, such as ‘Fear-of-happiness’ on Emotion Regulation, as well as ‘Independent self-construal (on Openness) and Interdependent self-construal’ (on Agreeableness). These constructs are prominent in cross-cultural research on individual differences but are rarely considered in the social-emotional learning literature, given that most of the educational frameworks for social-emotional skills were developed within single cultures.

5.3. Trait Building Blocks and Skills

Finally, the second column in Table 5.1 illustrates how most of these FFM qualities can be translated into skill constructs. Facets of the Emotional Stability domain can be transformed into emotion-regulation skills. For example, the Emotion-Regulation Skills Questionnaire (ERSQ; Berking & Znoj, 2008) describes a set of emotion regulation skills, including among others, emotion awareness, emotion acceptance, emotion reappraisal and emotion modification skills. Likewise, the lower-level Extraversion constructs of assertiveness, activity/enthusiasm, and sociability (initiating contact and connect to others) easily translate into assertiveness, presentation, and social contact skills, whereas these latter skills sets are necessary for developing leadership skills. The facets of Agreeableness form the heart of collaboration and communication skills, whereas the FFM facets orderliness, self-discipline, task engagement and achievement orientation culminate into task engagement and goal setting skills. Finally the domain of Openness to experience refers to qualities that have to do with being creative and open to innovation,

demonstrating eagerness to learn and curiosity, appreciating beauty and developing aesthetic sensitivity, being open and reflect on your own feelings, and having independent thought (autonomy). These trait qualities form the basis of creativity, appreciation, self-reflective and critical thinking skills.

The framework suggested in Table 5.1 further connects with other classification and taxonomic schemes developed for related areas, such as labour market and human resources competencies (see De Fruyt, Wille & John, 2015). The skills grouped in the 'Managing emotions' set are primarily intrapersonal competencies, whereas the skills related to extraversion and agreeableness are typically grouped in competency models under the header of interpersonal competencies. The agreeableness by conscientiousness skill area refers to skills that relate to Morality and Character, such as being reliable and respecting, including responsible decision-making and being accountable for decisions. Finally, the skills related to the openness by conscientiousness dimensions tap into Learning and Achieving competencies, reflected in the notion of life-long learning advocated in all educational frameworks nowadays.

6. VALIDATION STRATEGY

This section outlines a strategy to validate the social and emotional framework presented in this report. The validation process, which will take place during the feasibility study (2015-18) of the Longitudinal Study of Skill Development in Cities, involves the following major steps:

- Identification, adaptation and development of a range of appropriate measurement instruments;
- Measurement of a comprehensive set of children’s social and emotional skills across grades 1-12 in participating countries/cities; and
- Content, external, structural and cross-cultural validity tests.

The core underlying principles of the validation strategy are as follows:

Start with best existing measures: There is already a wealth of information on measurement instruments and their measurement properties for a range of social and emotional skills constructs. This study will build on the past research, and focus on areas where more work is needed to improve existing measures.

Employ multiple methods and scale adjustments: Given that none of the existing instruments provide precise measures of social and emotional skills, the optimal strategy is likely to include multiple sources of data and triangulate them to reduce measurement errors and biases (see also Annex 2). In addition to the core data sources, namely reports by self, teachers and parents, we suggest considering other types of measures, such as performance tests, behavioural indicators (e.g., truancy) and situational judgment tests. There is also a need to consider adjusting rating scales by using anchoring vignettes and forced-choice methods. This will help reduce various biases that plague rating scales such as cross-cultural differences, social desirability, reference group bias and response style bias.

Ensure cross-cultural relevance, comparability, and invariance: The measurement instruments must provide good measures of the latent socio-emotional constructs of interest. Moreover, the scaling of these instruments must also be comparable across participating countries/cities and population groups within a country/city. Cross-cultural invariance might be improved using performance tests, behavioral measures, anchoring vignettes and forced choice, though measurement equivalence will also need to be demonstrated for these methods.

Malleability and age relevance: Some may argue that social and emotional development can be highly age-dependent, and that some of the social and emotional skills may not have developed sufficiently before a child reaches certain ages. For instance we may not anticipate “interdependent self-construal” (e.g., I feel my fate is intertwined with the fate of those around me) to have developed among many children during the early grades. The feasibility study may provide relevant information on the ‘starting grade’ and ‘frequency’ in which each of the social and emotional skills may usefully be measured. The former can be assessed by evaluating from what age do children start understanding and developing a particular socio-emotional concept. The latter can be indirectly tested by assessing to what extent do social and emotional skills vary across different grades and children (after taking into account some of the individual differences such as demographic and socioeconomic background of children and parents).

Stakeholder consultation: It is of vital importance that the conceptual framework use labels of social and emotional skills (both at the facet and factor levels) that are well recognised by education stakeholders. After extensive psychometric analysis is performed, some of the labels used in the current framework could be validated or adjusted as necessary through consultation with superintendents, teachers, parents, employers and education officials.

7. POLICY QUESTIONS

The proposed Social and Emotional Skills Framework is a guiding principle for developing instruments to measure social and emotional skills. Together with the instruments to measure learning contexts (family, school and community) and outcomes (e.g. tertiary attainment, labour market and health), they will help address the following questions that are considered pertinent for policy-makers, teachers, school administrators and parents:

- **Which socio-emotional skills of children predict their cognitive, educational, labour market and social outcomes?**
- **Which family learning contexts predict children’s social and emotional development?**
 - Parenting activities (e.g., involvement in childcare and childrearing tasks and time spent with children)
 - Attitudes towards parenting (parenting stress, parental satisfaction)
 - Family resources (e.g., education, income, employment, benefit dependency)
- **Which school learning contexts predict children’s social and emotional development?**
 - School environment (e.g., classroom environment, violence)
 - Curricular and instructional practices (e.g., contents, delivery methods)
 - Teacher characteristics (e.g., age, gender, years of experience)
 - Extra-curricular activities (e.g., contents, objectives, delivery methods)
 - School resources (e.g., infrastructure, class materials, class size, child-to-staff ratio)
- **Which community learning contexts predict children’s social and emotional development?**
 - Community learning environment (e.g., availability of civic and cultural activities)
 - Community safety (e.g., violence, quality of life)
 - Community characteristics (e.g., peer’s socioeconomic background and values)
 - Community resources (e.g., public services, welfare regimes)
- **How malleable are social and emotional skills?** Are similar patterns observed across participating cities?
- **What are the social and emotional skills gaps by children’s gender and parental SES?**
- **To what extent do these skills gaps vary across cities?** Are cities with narrower skills gaps those that provide strong support to parents and children?
- **Do these skills gaps grow over time?**
- **Do children’s learning contexts prior to formal schooling (e.g., parenting, pre-school) predict children’s skill development during the first years of formal schooling,** even after accounting for families’ socio-economic background?

The effectiveness of addressing these questions will depend on the constructs and quality of instruments employed to measure learning contexts and outcomes. The developmental work on learning contexts and outcomes is scheduled to start in 2017. Moreover, the choice of empirical models used to identify the dynamics of social and emotional skill formation will also affect the precision in which these questions can be addressed. The latest OECD report: *‘Skills for Social Progress: The powers of social and emotional skills’* (OECD, 2015), presents an innovative method that takes into account: (a) skills today depends on the previous skills and investment in skills made during the previous year, (b) previous skills affect the productivity of mobilising new investments in skills, (c) previous socio-emotional skills affect the development of both socio-emotional and cognitive skills, (d) independent effects as well as interactive effects of skills (e.g. the impact of literacy on depression is larger the higher the self-esteem) on education, labour market and social outcomes. The longitudinal data structure and the range of measurement instruments to be proposed in this study will permit application of such elaborate empirical models in order to better shed light on the key questions for policy-makers and educators.

ANNEX A. ALPHABETICAL LIST OF 21ST CENTURY CHARACTERISTICS

	Diligence	Integrity	Reverence
Ability to quickly acquire and apply new knowledge			
Abnegation	Discipline	Interconnectedness	Risk taking
Abstract problem solving	Diversity	Interdependency	Self-actualization
Acceptance	Efficiency	Justice	Self-awareness
Accountability	Effort	Kindness	Self-care
Adaptability	Empathy	Leadership	Self-compassion
Altruism	Energy	Leading by example	Self-control at school
Applying technology	Engagement	Learning from mistakes and failures	Self-control in relationships
Appreciation	Enthusiasm	Listening to others	Self-direction
Appreciating beauty in the world	Equanimity	Living in harmony with nature	Self-discipline
Appreciating others	Equity	Living in harmony with others	Self-esteem
Appreciating what I have	Ethics	Load management	Self-kindness
Assertiveness	Excitement of creating something new	Love	Self-reflection
Authenticity	Executing plans, follow through	Loyalty	Self-respect
Balance	Existentiality	Mental flexibility	Selflessness
Belonging	Exploration	Mentorship	Sensibility
Benevolence	Fairness	Mercy	Sharing
Bravery	Feedback	Mindfulness	Social awareness
Camaraderie	Feeling awe	Modesty	Social intelligence
Care	Flexibility	Motivation	Social perspective
Charisma	Focus	Negotiation	Socialization
Charity	Followership	Observation	Speaking out, taking a stand
Cheerfulness	Following	Oneness	Spirituality
Citizenship	Forgiveness	Open-mindedness	Spontaneity
Civic-mindedness	Fortitude	Optimism	Sportsmanship
Commitment	Generosity	Organization	Spunk
Common humanity	Genuineness	Passion	Stability
Compassion	Goal orientation	Patience	Tackling tough problems
Confidence	Grace	Perseverance	Teamwork
Conscientiousness	Gratitude	Persistence	Tenacity
Consciousness	Grit	Playfulness	Timeliness
Consideration	Growth	Precision	Tinkering / inventing
Consistency	Happiness	Presence	Tolerance
Cooperation	Helpfulness	Problem solving	Toughness
Courage	Heroism	Productivity	Tranquility
Critical thinking	Honesty	Professionalism	Trustworthiness
Cross-cultural awareness	Honor	Project management	Truthfulness
Curiosity	Humaneness	Prudence	Verve
Dealing with ambiguity	Humbleness / humility	Public speaking	Vigor
Decency	Humor	Receptivity	Virtue
Decisiveness	Inclusiveness	Reliability	Vision
Decorum	Initiative	Resilience	Willingness to try new ideas
Delegation	Innovation	Resourcefulness	Wonder
Dependability	Inquisitiveness	Respect for others	Work ethic
Determination	Insight	Responsibility	Zeal
Devotion	Inspiration	Results orientation	Zest

Note. Items derived from Trilling & Fadel (2009) and Fadel (March 2014). Table adapted from John & Mauskopf (2014).

ANNEX B. FEASIBILITY STUDIES: INITIAL METHODOLOGICAL CONSIDERATIONS

We have discussed various issues related to assessment and research design throughout this report, including the need to supplement self-report data with information collected from other data sources, like parents, teachers, or potentially coaches (e.g., in the US high school system, they tend to know the kids involved in athletics better than most teachers do), as well as data from school records (e.g., lateness) and possibly situational judgment tests, like the MSCEIT. In this final section, we offer some initial thoughts to stimulate consideration and further discussion by the expert group.

One critical set of more technical issues involves the minimum sample size required to test and validate the items to be used for self-reports, as well as for teacher and parent ratings. Other issues involve the grade levels that should be studied and what kind and how much data needs to be obtained from each student, at multiple occasions, at what grade levels children can start providing meaningful self-reports with acceptable levels of internal consistency and differentiation among the concepts rated, and how many items children at particular ages can be expected to complete in any one testing session.

Age ranges for self-reports by children and adolescents

We have commented earlier that adults have tended to underestimate the capacity of children to provide reliable personality ratings. Using a puppet interview technique, Measelle, John, Ablow, Cowan, and Cowan (2005) found that children as young as age 5-7 were able to report on their emerging self-concept in terms of the Big Five domains, with modest but significant internal consistency, temporal stability, discriminant validity among the Big Five domains, and external validity as assessed with teacher and parent ratings. An individually administered puppet interview is certainly not a feasible assessment device for the proposed OECD study. However, one could conceivably develop a computer-based measure consisting of 50 short animations showing relevant behaviours or emotions that are inspired by the content of the puppet interview and use a similar “is like me” or “is not like me” response format.

In terms of late childhood and early adolescence, age 10 has been the youngest age studied using existing Big Five measures. Using the BFI, Soto, John, Gosling, and Potter (2008) found that in the US, by age 14, self-reports by adolescents were essentially indistinguishable from those of adults; moreover, when problems with acquiescence and rating scale usage were controlled, many kids as young as age 10 could provide self-reports with reasonable psychometric characteristics, such as internal consistency and discriminant validity. The BFI was designed to have a 5th grade reading level; the starting ages for the slightly more difficult NEO PI-R items tend to be closer to age 12.

Moreover, the very large and economically diverse data set collected in schools by Primi and colleagues (2014, submitted) in Brazil generally replicated the US findings; when variation in rating scale usage was controlled, the Big Five factors could be identified reliably in self-reports of children as young as age 10-12. However, some caveats apply. The Brazilian self-report items had been carefully tailored and pilot-tested to conform to the language used by these children, and the number of items given in one assessment session had been limited to a maximum of 65 short items for the younger groups (ages 10-13) and 95 for the older age groups (14-18). Moreover, these findings apply only to measures of the broad Big Five domains; we know much less about facet level traits. Future development work needs to examine whether discriminations among facets within Big Five domains can be made reliably before age 14; for example, Soto et al.’s (2008) analyses suggest that the facets of extraversion may not cohere into one unified domain as early as agreeableness and conscientiousness.

Sample size considerations

In general, research suggests that sample sizes of N=500 will generate stable factor structures and reliability estimates in kids' self-reports for up to 100 items. Thus, one might suggest minimum samples of 500 girls and 500 boys, to analyse them separately and check that results generalize, so total N=1000. This should also be enough to examine effects of SES (lowest 20% = 100 boys and 100 girls, vs. the rest) and test the social-perceptual and language skills needed to make judgments about self and other (e.g., as assessed by vignettes requiring children to rate hypothetical others on the same items). These sample size estimates assume there are no critical minority groups, for whom separate estimates or analyses have to be run. If there are, larger samples would have to be drawn.

It will be critical to ensure that the items are easy enough to understand and do apply to kids in grade 7 (expected ages 12-13) but are not too childish and thus still useful for kids up to grade 12 (age 17-19). One design to test for measurement equivalence at reasonable age intervals would be to assess 500 boys and 500 girls in grade 7, another in grade 9, and another in grade 11. This would also permit an initial test of the expected plasticity effects, where age differences should be apparent at least in the facets related to agreeableness and conscientiousness (and gender differences should appear and increase in magnitude in the negative-affect facets).

Number and type of items to be completed

If previous research is any guide, the 11th and 12th graders will be able to respond to 100 BFI-type items in about 20-30 minutes (depending on their prior experience with rating tasks), thus leaving time to administer vignettes to obtain ratings of others. So, if schools can provide 60 min of assessment time, that would permit the administration of an initial unrefined item pool of 200 items plus one vignette set rating for each of the Big Five domains (but not the for each of the facet constructs). However, if the starting point is 300 initial candidate items to select the final (let's say 90) items, either more time (or multiple sessions) will be needed, or more participants are needed to permit a partial-administration design.

These estimates are likely to hold with non-minority students in developed countries. In Brazil, our colleagues Primi and Santos (see Primi et al., 2014) concluded from their extensive pilot testing that they should *not give more than 100 items per session even to 12 graders*, given the particular limitations in reading, concentration, and classroom set-up. This becomes a bigger issue in the younger groups, with 40 - 60 items doable for ages 10-11; if starting with age 12 in 7th grade, K=100 items may be feasible per session but one can expect more than 100 items to try out in the Round 1 of measure development. Thus, it may be necessary to double or triple the above estimates of sample sizes, so that different kids can do smaller but different subsets of the items.

Applicable for multi-source ratings

Paraphrasing Wim Hofstee (1994) in the title of this section, a key question is: "Who is the best informant on children's developing social-emotional skills"? Most approaches to personality assessment agree that appraisals of an individual must be based on multi-informant designs, rather than a single source of data (Wiggins, 1973). In the present context, multiple different stakeholders, such as the children and adolescents themselves as well as their parents and their teachers, can provide unique perspectives on children's standing and development of skills, in addition to common variance.

The evaluation of social-emotional skills is usually done via self- or observer ratings on how children and adolescents typically behave or how well they can do so in general or in more specific situations. Alternatively, more situational judgment type of assessments, where children are presented a situation description and have to select the most appropriate response from a predefined set of reactions,

are used infrequently and have their own shortcomings. Major problems are that they usually span a very small range of constructs relative to the number of items and time necessary to measure a construct reliably. Besides, these methods identify if people are able to select that particular answer for which subject matter experts agree that this is the most appropriate response, which is very different from explaining how people will react in daily life.

Children can provide reliable and valid self-descriptions on personality and social-emotional skill descriptive items on average starting from the age of 10 onwards (Soto, John, Gosling, and Potter, 2011). This capacity is dependent on a series of critical factors, including language proficiency, but also cognitive and social development. First, children need to have acquired a certain vocabulary and a basic reading level to be in a position to administer the assessment. Simplicity and clarity in language is anyway an important requirement for skill descriptive items, because assessments not only have to be completed by children and adolescents themselves, but often also by parents of different socio-economic classes, and teachers who will have to rate multiple children in their class. These constraints require grammatically streamlined and short items, an easily understandable response scale format, and clear instructions. Many of these principles have been used when constructing the Big Five Inventory (BFI; John, Donahue, and Kentle, 1991) or the Hierarchical Personality Inventory for Children (HiPIC; Mervielde and De Fruyt, 2002). Explicit guidelines for the item-writing process are provided by Hendriks, Hofstee and de Raad (Hendriks, Hofstee, and De Raad, 1999).

Probably more important is that children also need to have developed some first self-reflective and social-comparison skills. According to Barenboim (1981), children first make behavioural comparisons (e.g. “Ricardo runs faster than Daniel”), and start to actively use trait terms thereafter (e.g. “Trudy is shy”). Furthermore, children’s person perception skills need to develop into a multidimensional scheme, to a point where they have a notion that multiple independent trait attributes may apply to themselves or another person. During development, children first associate a single individual with one characteristic [see for example the figures portrayed in children’s books and comics that are even named after a single trait, e.g. the different smurfs, each with their typifying characteristics, Asterix (small but smart) and Obelix (raw power), gnome “Lui” (lazy), ...], and this perspective needs to progress into a multidimensional space of person-perception that can be used to describe differences between, but also within persons. The evidence available right now suggests that this is achieved by age 10-11, in line with the emergence of formal-operational thinking.

Besides self-reports, observer reports of skills and personality are also frequently used and form a necessary amendment to self-descriptions. For children in primary school, parents and teachers often act as first informants, whereas in secondary school parents and teachers often complement adolescent self-ratings. Research shows that all perspectives have unique and valuable viewpoints on individual differences, with father and mother ratings correlating around .60 to .70, teachers with parents correlating between .30 and .60, and parent and teacher ratings correlating around .30 with children/adolescents’ self-ratings (Mervielde and De Fruyt, 2009). The magnitude of these correlations suggests that all perspectives share some variance, but also have their own specific and informative viewpoint. Teachers, for example, have experience with children in the more structured context of the classroom, and are in a good position to observe more interpersonal and task-oriented skills, whereas parents provide ratings relying on the home-context. In addition, teachers rely on a much broader frame-of-reference to describe pupils’ characteristics, because they accumulate professional experience with 20 new pupils in their classroom each year, whereas the scope of parents is usually much smaller and more idiosyncratic. De Los Reyes and colleagues (2013), for example, suggested that parents and teachers may have different “decision thresholds” for concluding that behaviour is problematic. Finally, Rescorla et al. (2014) have conducted one of the most impressive studies on parent-teacher agreement on children’s problem behaviours across 21 societies. There were striking similarities across societies: parents reported higher problem scores than teachers across societies (with some differences in magnitude across cultures), and similar age and gender differences were

observed. Rescorla et al. (2014) also found, that within and across societies, parents and teachers agreed strongly on items that received low, medium, and high ratings.

In sum, the overall recommendation across many studies is to advocate cross-informant assessment because it brings shared but also complementary perspectives to the study of personality, social-emotional skills or problem behaviours. Informant discrepancies should hence not be considered by default as measurement error, though these viewpoints bring substance to the discussion. In addition, a multi-informant design has many psychometric advantages, including better ways to deal with common method variance to explain criteria and enhanced possibilities to estimate the variables at stake.

Ease of administration

Finally, constructs and accompanying measures will have to be relatively easy to administer to all informants (via an electronic assessment platform). Assessments will have to be completed by children and adolescents themselves, parents of different socio-economic classes, and teachers who will have to rate multiple children in their class. These constraints require grammatically streamlined and short items, an easily understandable format for the response scale to be used, and clear instructions. Guidance for such an approach can be found in several examples, such as the Big Five Inventory (BFI; John, Donahue, & Kentle, 1991) or the Hierarchical Personality Inventory for Children (HiPIC; Mervielde & De Fruyt, 2002).

Assessing short-term “transient error”: Need for multiple sessions

Factor structure and reliability involve internal consistency across the test items, which can be assessed within a single testing occasion (or time), so those basic issues can be addressed in the design sketched out so far. However, another critical unknown when aiming to assess *change* is what has been called “transient error” (Chmielewski and Watson, 2009)—that is, low temporal reliability of the items within the same individual over different occasions (e.g., a week or a month at most). If that temporal error variation is too high, and we don't know that in advance, we will not be able to separate true, lasting, long-term change within the individual (e.g., change over 2 years) from short-term error variations within the same individual. Existing studies and data sets do not have data on kids, especially for comparing grades 7 and 12, and even adult data are very rare because repeated measurements are required. But kids are expected to attend school 5 days per week, thus it will be possible for the OECD study to obtain these kinds of data. So, transient error is a critical issue to examine in the feasibility studies. For example, the data may show that *more items* are needed in grade 7 to reach acceptable levels of short-term stability than in the later grades.

Context effects on measurement

Likely of lesser importance than transient error is that researchers do not know much about effects of context in schools on assessments, such as day of the week (e.g., Monday vs. Friday?) and time of day (e.g., early morning vs. before lunch vs. after lunch?). One might think that such effects are small but if they do not get examined in the feasibility studies, then context effects may have to be held constant in the longitudinal study itself (e.g., all assessment have to be completed before lunch). In longitudinal college studies, it can make a big difference whether assessments are done in the beginning or end of the semester, or close to finals, etc. Similar considerations would seem to apply here.

Teacher and parent ratings

Similar considerations apply to the younger cohorts in grades 1 to 6, for teacher and parent ratings. However, as these ratings are made from an external perspective, they tend to be less differentiated (more internally consistent across items) and less variable (more stable) over time. So, especially teacher's ratings can be obtained using fewer items, and assessing “transient error” (or short-term temporal stability) could

be done in a subset of the teachers providing ratings (50% or even only 33% of the teachers or parents would have to be asked to provide their ratings a second time).

Mapping the anticipated measurement transition from grades 6 to 7

The issue of mapping the different data sources (e.g., teacher ratings in grade 6 and self-reports in grade 7) onto each other should be addressed by obtaining both data sources in either grade 6 or 7 (or preferably in both). That is, ideally the entire pilot sample from grades 6 and 7 would provide teacher or parent ratings as well as self-reports. This information will be critical in making sure that one can developmentally map the information collected from grades 1 to 6 with the information collected from grades 7 to 12. Again, the convergence of the 3 data sources ought to be tested separately for boys and girls, and possibly for ethnic minority groups, or even for high vs. low SES subgroups.

A related issue is how many "age specific" or "age appropriate" measures would have to be introduced. There is reason to assume that the same measure could be used from Grade 7 to 12. Recall that Soto et al. (2008, 2011) used the same BFI items (with their 5th grade readings levels) for ages 10 to 18. However, it would seem likely that the younger kids in grades 5 and 6 and maybe 7 (ages 10-13) might do better with a measure specifically designed for their age group, and that the more advanced conceptual language appropriate for older kids might best be limited to the high school grades (9-12). The feasibility studies will have to explore those possibilities.

REFERENCES

- Allport, G. W. (1958). What units shall we employ? In G. Lindzey (Ed.), *Assessment of human motives* (pp.238-260). New York: Rinehart.
- Ananiadou, K., and M. Claro (2009). 21st century skills and competencies for New Millenium Learners in OECD countries. Paris, France: Centre for Educational Research and Innovation (CERI) - New Millenium Learners.
- Barenboim, C. (1981). The development of person perception in childhood and adolescence - from behavioral-comparisons to psychological constructs to psychological comparisons. *Child Development*, 52(1), 129-144. doi: 10.2307/1129222
- Benet-Martinez, V., and S. Oishi (2008). Culture and personality. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 542-567). New York: Guilford.
- Benet-Martinez, V., and O. P. John. (1998). Los Cinco Grandes across cultures and ethnic groups: Multitrait multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology*, 75(3), 729-750. doi: 10.1037/0022-3514.75.3.729
- Berking, M., and H. Znoj (2008). Entwicklung und Validierung eines Fragebogens zur standardisierten Selbsteinschaätzung emotionaler Kompetenzen [Development and validation of a self-report measure for the assessment of emotion-regulation skills]. *Zeitschrift fur Psychiatrie, Psychologie und Psychotherapie*, 56, 141–152.
- Bipp, T., and K. Van Dam (2014). Extending hierarchical achievement motivation models: The role of motivational needs for achievement goals and academic performance. *Personality and Individual Differences*, 64, 157-162. doi: 10.1016/j.paid.2014.02.039
- Blair, C. (2002). School readiness. Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*, 57 (2), 111-127.
- Briley, D. A., Domiteaux, M., and E. M. Tucker-Drob (2014). Achievement-relevant personality: Relations with the Big Five and validation of an efficient instrument. *Learning and Individual Differences*, 32, 26-39.
- Catanzaro, S. J., and J. Mearns (1990). Measuring generalized expectancies for negative mood regulation: Initial scale development and implications. *Journal of Personality Assessment*, 54, 546-563.
- CCR (2015). Character education for a challenging century: What should students learn? Downloaded from www.curriculumredesign.org
- Cheung, F. M. et al. (2001). Indigenous Chinese personality constructs - Is the five-factor model complete? *Journal of Cross-Cultural Psychology*, 32(4), 407-433. doi: 10.1177/0022022101032004003
- Cheung, F. M. et al. (1996). Development of the Chinese Personality Assessment Inventory. *Journal of Cross-Cultural Psychology*(27), 181-199.

- Chmielewski, M., and D. Watson. (2009). What is being assessed and why it matters: The impact of transient error on trait research. *Journal of Personality and Social Psychology*, 97(1), 186-202. doi: 10.1037/a0015618
- Church, A. T. (2010). Measurement issues in cross-cultural research. In G. Walford, E. Tucker and M. Viswanathan (Eds.), *The SAGE handbook of measurement* (pp. 151-177). Los Angeles: Sage.
- Church, A. T. et al. (2011). Are Cross-Cultural Comparisons of Personality Profiles Meaningful? Differential Item and Facet Functioning in the Revised NEO Personality Inventory. *Journal of Personality and Social Psychology*, 101(5), 1068-1089. doi: 10.1037/a0025290
- Connelly, B. S., and D. S. Ones (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, 136(6), 1092-1122. doi: 10.1037/a0021212
- Costa, P. T., and R. R. McCrae, (1992). *Revised NEO Personality Inventory and Five-Factor Inventory Professional Manual*. Odessa, FL: Psychological Assessment Resources.
- Coutinho, S. A. (2007). The relationship between goals, metacognition, and academic success. *Educate*, 7(1), 39-47.
- De Bolle, M. et al. (2012). General personality and psychopathology in referred and nonreferred children and adolescents: An investigation of continuity, pathoplasty, and complication models. *Journal of Abnormal Psychology*, 121(4), 958-970. doi: 10.1037/a0027742
- Deci, E. and R. Ryan (2002). *Handbook of Self-Determination Research*. Rochester, NY: The University of Rochester Press.
- De Fruyt, F. et al. (2009). Assessing the universal structure of personality in early adolescence with The NEO-PI-R and NEO-PI-3 in 24 Cultures. *Assessment*, 16(3), 301-311. doi: 10.1177/1073191109333760
- De Fruyt, F. et al. (2006). Police interview competencies: Assessment and associated traits. *European Journal of Personality*, 20(7), 567-584.
- De Fruyt, F., et al. (2006). Five types of personality continuity in childhood and adolescence. *Journal of Personality and Social Psychology*, 91(3), 538-552.
- De Fruyt, F., and B. De Clercq. (2014). Antecedents of personality disorder in childhood and adolescence: Toward an integrative developmental model. *Annual Review of Clinical Psychology*, 2014(10), 449-476.
- De Fruyt, F., and J. P. Rolland (2013). *Personality for Professionals Inventory: PfpI - Handleiding*. Amsterdam: Pearson.
- De Fruyt, F., and K. Van Leeuwen (2014). Advancements in the field of personality development. [Article]. *Journal of Adolescence*, 37(5), 763-769. doi: 10.1016/j.adolescence.2014.04.009
- De Fruyt, F., and B. Wille (2013). Cross-cultural issues in personality assessment. In N. D. and Tett Christiansen, R. P. (Ed.), *Handbook of Personality at Work* (pp. 333-354). New York.

- De Fruyt, F., Wille, B., and O. P. John, (2015). Employability in the 21st century: Complex (interactive) problem solving and other essential skills. *Industrial and Organizational Psychology*, 8, 2, 276-281.
- De Haan, A. D. et al. (2013). Developmental personality types from childhood to adolescence: Associations with parenting and adjustment. *Child Development*, 84(6), 2015-2030. doi: 10.1111/cdev.12092
- De Los Reyes, A. et al. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*, Vol 9, 9, 123-149. doi: 10.1146/annurev-clinpsy-050212-185617
- Debusscher, J., Hofmans, J., and F. De Fruyt (2014). The curvilinear relationship between state neuroticism and momentary task performance. *Plos One*, 9(9). doi: e10698910.1371/journal.pone.0106989
- Decuyper, M. et al. (2013). Latent personality profiles and the relations with psychopathology and psychopathic traits in detained adolescents. *Child Psychiatry and Human Development*, 44(2), 217-232. doi: 10.1007/s10578-012-0320-3
- Decuyper, M. et al. (2009). A meta-analysis of psychopathy-, antisocial PD- and FFM associations. *European Journal of Personality*, 23(7), 531-565. doi: 10.1002/per.729
- Denissen, J. J. A. (2014). A Roadmap for further progress in research on personality development. *European Journal of Personality*, 28, 213-215.
- DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology*, 91, 6, 1138-1151.
- Diener, E. (2013). The remarkable changes in the science of subjective well-being. *Perspectives on Psychological Science*, 8(6), 663-666. doi: 10.1177/1745691613507583
- Diseth, A. (2003a). Personality and approaches to learning as predictors of academic achievement. [Article]. *European Journal of Personality*, 17(2), 143-155. doi: 10.1002/per.469
- Diseth, A. (2013b). Personality as an indirect predictor of academic achievement via student course experience and approach to learning. *Social Behavior and Personality*, 41(8), 1297-1308. doi: 10.2224/sbp.2013.41.8.1297
- Duckworth, A. L., and M. E. P. Seligman (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, 98(1), 198-208.
- Duckworth, A. L., et al. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), 1087-1101. doi: 10.1037/0022-3514.92.6.1087
- Durlak, J. A. et al. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405-432. doi: 10.1111/j.1467-8624.2010.01564.x
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. New York: Psychology Press.
- Elias, M. J. et al. (Eds.). (1997). *Promoting social and emotional learning: Guidelines for educators*. Alexandria, VA: Association for Supervision and Curriculum Development.

- English, T. and O. P. John (2013). Understanding the social effects of emotion regulation: The mediating role of authenticity for individual differences in suppression. *Emotion*, 13, 314-329.
- Entwistle, N. J., and H. Tait (1995). *The revised approaches to studying inventory*. Edinburgh: Centre for Research on Learning and Instruction.
- Epstein, M. H. (2004). *Behavioral and Emotional Rating Scale-2: A strength-based approach to assessment*. Austin, TX: PRO-ED.
- Erikson, E. H. (Ed.). (1950). *Childhood and society*. New York: Norton.
- Fanous, A. H. et al. (2007). A longitudinal study of personality and major depression in a population-based sample of male twins. *Psychological Medicine*, 37(8), 1163-1172. doi: 10.1017/s0033291707000244
- Fischer, S., Smith, G. T., and M. A. Cyders (2008). Another look at impulsivity: A meta-analytic review comparing specific dispositions to rash action in their relationship to bulimic symptoms. *Clinical Psychology Review*, 28(8), 1413-1425. doi: 10.1016/j.cpr.2008.09.001
- Flavell, J. H. (1979). Meta-cognition and cognitive monitoring - new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906-911. doi: 10.1037/0003-066x.34.10.906
- Føllesdal, H., and K. A. Hagtvet (2009). Emotional intelligence: The MSCEIT from the perspective of generalizability theory. *Intelligence*, 37, 94-105.
- Friedman, H. S., and M. L. Kern (2014). Personality, Well-Being, and Health. In S. T. Fiske (Ed.), *Annual Review of Psychology*, Vol 65 (Vol. 65, pp. 719-742). Palo Alto: Annual Reviews.
- Furlong, M.J., et al. (2007). Cross-Validation of the Behavioral and Emotional Rating Scale-2 Youth Version: An Exploration of Strength-Based Latent Traits. *Journal of Child and Family Studies*. DOI 10.1007/s10826-006-9117-y
- Furnham, A. (1996). The FIRO-B, the learning style questionnaire, and the five-factor model. *Journal of Social Behavior and Personality*, 11(2), 285-299.
- Furnham, A., and J. Taylor (2004). *The dark side of behaviour at work: Understanding and avoiding employees leaving, thieving and deceiving*. Hampshire: Palgrave MacMillan.
- George, L.G., Helson, R., and O. P. John (2011). The “CEO” of women's work lives: How Big Five Conscientiousness, Extraversion, and Openness predict 50 years of work experiences in a changing sociocultural context. *Journal of Personality and Social Psychology*, 101, 812-830.
- Gerlach, G., Herpertz, S., and S. Loeber (in press). Personality traits and obesity: a systematic review. *Obesity Review*. doi: doi: 10.1111/obr.12235
- Goldberg, L. R. et al. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40(1), 84-96.
- Goleman, D. (1995). *Emotional Intelligence: Why It Can Matter More Than IQ*, New York, Bantam
- Gratz, K. L., and L. Roemer (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment*, 26, 41-54.

- Halverson, C. F. et al. (2003). Personality structure as derived from parental ratings of free descriptions of children: The inventory of child individual differences. *Journal of Personality*, 71(6), 995-1026. doi: 10.1111/1467-6494.7106005
- Hampson, S. E., John, O. P., and L. R. Goldberg (1986). Category breadth and hierarchical structure in personality: Studies of asymmetries in judgments of trait implications. *Journal of Personality and Social Psychology*, 51, 37-54.
- Hawes, D. J., Price, M. J., and M. R. Dadds (2014). Callous-unemotional traits and the treatment of conduct problems in childhood and adolescence: A comprehensive review. *Clinical Child and Family Psychology Review*, 17(3), 248-267. doi: 10.1007/s10567-014-0167-1
- Helson, R. et al. (2002). The growing evidence for personality change in adulthood: Findings from research with personality inventories. *Journal of Research in Personality*, 36, 287-306.
- Hendriks, A. A. J., Hofstee, W. K. B., and B. De Raad (1999). The Five-Factor Personality Inventory (FFPI). *Personality and Individual Differences*, 27(2), 307-325.
- Hills, P., and M. Argyle (2001). Emotional stability as a major dimension of happiness. *Personality and Individual Differences*, 31(8), 1357-1364. doi: 10.1016/s0191-8869(00)00229-4
- Hoekstra, H. A., and E. Van Sluijs (Eds.). (2003). *Managing competencies: Implementing human resource management*. Nijmegen: Royal Van Gorcum.
- Hofstee, W. K. B. (1994). Who should own the definition of personality? *European Journal of Personality*, 8(3), 149-162.
- Hofstee, W. K. B., Deraad, B., and L. R. Goldberg (1992). Integration of the Big-5 and Circumplex Approaches to Trait Structure. *Journal of Personality and Social Psychology*, 63(1), 146-163.
- Hogan, J., and B. Holland (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology*, 88(1), 100-112.
- Honey, P., and A. Mumford (1982). *The Manuals of Learning Styles*. Maidenhead: Honey Press.
- Hurtz, G. M., and J. J. Donovan (2000). Personality and job performance: The big five revisited. *Journal of Applied Psychology*, 85(6), 869-879.
- John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66-100). New York: Guilford Press.
- John, O. P. et al. (1994). The "Little Five": Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Development*, 65, 160-178.
- John, O. P., and J. Eng (2014). Three approaches to individual differences in affect regulation: Conceptualization, measures, and findings. In J.J. Gross (Ed.), *Handbook of emotion regulation* (2nd ed.). New York: Guilford.
- John, O. P., and J. J. Gross (2007). Individual differences in emotion regulation. In J.J. Gross (Ed.), *Handbook of emotion regulation* (pp. 351-372). New York: Guilford.

- John, O. P., and J. J. Gross (2004). Healthy and unhealthy emotion regulation: Personality processes, individual differences, and life span development. *Journal of Personality*, 72, 1301-1333.
- John, O. P., and S. Mauskopf (2015, June 4). Self-reported socio-emotional qualities: Five factors for 21st century skills? Poster presented at the Bi-annual Meetings of the Association for Personality Research, Saint Louis, Missouri, USA.
- John, O. P., and S. Srivastava (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin, and O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102-138). New York: Guilford.
- John, O. P., Donahue, E. M., and R. L. Kentle (1991). The "Big Five" Inventory - Versions 4a and 54. Berkeley: University of California, Berkeley, Institute of Personality and Social Research.
- John, O. P., Naumann, L., and C. J. Soto (2008). Paradigm shift to the integrative Big Five trait taxonomy: Discovery, measurement, and conceptual issues. In O.P. John, R.W. Robins, and L.A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114-158). New York: Guilford.
- Joseph, D.L., and D. A. Newman (2010). Emotional intelligence: An integrative meta-analysis and cascading model. *Journal of Applied Psychology*, 95, 54-78.
- Joshanloo, M. et al. (2014). Cross-cultural validation of fear of happiness scale across 14 national groups. *Journal of Cross-Cultural Psychology*, 45(2), 246-264. doi: 10.1177/0022022113505357
- Judge, T. A. et al. (1999). The Big Five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, 52(3), 621-652.
- Kautz, T. et al. (2014), "Fostering and Measuring Skills: Improving Cognitive and Non-cognitive Skills to Promote Lifetime Success", *OECD Education Working Papers*, No. 110, OECD Publishing, <http://dx.doi.org/10.1787/5jxsr7vr78f7-en>.
- Kern, M. L., and H. S. Friedman (2008). Do conscientious individuals live longer? A quantitative review. *Health Psychology*, 27(5), 505-512. doi: 10.1037/0278-6133.27.5.505
- Klimstra, T. A. et al. (2014). Personality and externalizing behavior in the transition to young adulthood: the additive value of personality facets. *Social Psychiatry and Psychiatric Epidemiology*, 49(8), 1319-1333. doi: 10.1007/s00127-014-0827-y
- Kohnstamm, G. A. et al. (1998). *Parental descriptions of child personality. Developmental antecedents of the Big Five?*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Kolb, D.A. (1984). *Experimental Learning*. Englewood Cliffs, New York: Prentice-Hall.
- Kolb, D.A. (1976). *Learning Style Inventory: Technical Manual*. Boston: MA: McBer.
- Kolko, David J. et al. (2009). Community vs. clinic-based modular treatment of children with early-onset ODD or CD: A clinical trial with 3-year follow-up. *Journal of Abnormal Child Psychology*, 37(5), 591-609. doi: 10.1007/s10802-009-9303-7
- Komarraju, M. et al. (2011). The Big Five personality traits, learning styles, and academic achievement. *Personality and Individual Differences*, 51(4), 472-477. doi: 10.1016/j.paid.2011.04.019

- Kotov, R. et al. (2010). Linking "Big" personality traits to anxiety, depressive, and substance use disorders: A meta-analysis. *Psychological Bulletin*, 136(5), 768-821. doi: 10.1037/a0020327
- Kwan, V. S. Y., Bond, M. H., and T.M. Singelis (1997). Pancultural explanations for life satisfaction: Adding relationship harmony to self-esteem. *Journal of Personality and Social Psychology*, 73(5), 1038-1051. doi: 10.1037/0022-3514.73.5.1038
- Kyllonen, P. C., and J. P. Bertling (2013). Innovative Questionnaire Assessment Methods to Increase Cross-Country Comparability. *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*, 277.
- Laceulle, O. M. et al. (2014). A test of the vulnerability model: temperament and temperament change as predictors of future mental disorders - the TRAILS study. *Journal of Child Psychology and Psychiatry*, 55(3), 227-236.
- Landy, F. (2005). Some historical and scientific issues related to research on emotional intelligence. *Journal of Organizational Behavior*, 26, 411-424.
- Lievens, F., and P. R. Sackett (2012). The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *Journal of Applied Psychology*, 97(2), 460-468. doi: 10.1037/a0025741
- Lievens, F., De Corte, W., and E. Schollaert (2008). A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology*, 93(2), 268-279. doi: 10.1037/0021-9010.93.2.268
- Lochman, J. E. et al. (2014). Does a booster intervention augment the preventive effects of an abbreviated version of the coping power program for aggressive children? *Journal of Abnormal Child Psychology*, 42(3), 367-381. doi: 10.1007/s10802-013-9727-y
- Loehlin, J. C. et al. (1998). Heritabilities of common and measure-specific components of the Big Five personality factors. *Journal of Research in Personality*, 32, 431-453.
- Lopes, P. N. et al. (2005). Emotion regulation ability and the quality of social interaction. *Emotion*, 5, 113 - 118.
- Markus, H. R., and S. Kitayama (1991). Culture and the self - implications for cognition, emotion, and motivation. *Psychological Review*, 98(2), 224-253. doi: 10.1037//0033-295x.98.2.224
- Martel, M. M. et al. (2014). Longitudinal prediction of the one-year course of preschool ADHD symptoms: Implications for models of temperament-ADHD associations. *Personality and Individual Differences*, 64, 58-61. doi: 10.1016/j.paid.2014.02.018
- Martel, M. M. et al. (2010). The structure of childhood disruptive behaviors. *Psychological Assessment*, 22(4), 816-826. doi: 10.1037/a0020975
- Mayer, J. D., and Y. N. Gaschke (1988). The experience and meta-experience of mood. *Journal of Personality and Social Psychology*, 55, 102-111.
- Mayer, J. D., Salovey, P., and D. R. Caruso (2002). *MSCEIT users manual*. MHS: Toronto.

- McCrae, R. R., and P. T. Costa (1996). Toward a new generation of personality inventories: Theoretical contexts for the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 51-87). New York/London: Guilford Press.
- McCrae, R. R., and O. P. John (1992). An introduction to the Five-Factor Model and its applications. *Journal of Personality*, 60, 175-215.
- McCrae, R. R., and A. Terracciano (2008). The Five-Factor Model and its correlates in individuals and cultures. In F.J.R. Van de Vijver, D. A. Van Hemert and Y. Poortinga (Eds.), *Multilevel analysis of individuals and cultures* (pp. 249-283). Mahwah, NJ: Erlbaum.
- McCrae, R. R., and A. Terracciano (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology*, 88(3), 547-561.
- Measelle, J. R. et al. (2005). Can children provide coherent, stable, and valid self-reports on the Big Five dimensions? A longitudinal study from ages 5 to 7. *Journal of Personality and Social Psychology*, 89, 90-106.
- Mervielde, I., and F. De Fruyt (2002). Assessing children's traits with the Hierarchical Personality Inventory for Children. In B. De Raad and M. Perugini (Eds.), *Big Five Assessment* (pp. 129-146). Gottingen: Hogrefe and Huber Publishers.
- Mervielde, I., and F. De Fruyt (1999). Construction of the Hierarchical Personality Inventory for Children (HiPIC). In I. Mervielde, I. Deary, F. De Fruyt and 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.
- Mervielde, I., De Fruyt, F., and B. De Clercq (2009). *Hiërarchische Persoonlijkheidsvragenlijst voor Kinderen [Hierarchical Personality Inventory for Children]: Handleiding*. Amsterdam: Hogrefe Publishers.
- Milfont, T. L., and C. G. Sibley (2012). The big five personality traits and environmental engagement: Associations at the individual and societal level. *Journal of Environmental Psychology*, 32(2), 187-195. doi: 10.1016/j.jenvp.2011.12.006
- Minbashian, A., and D. Luppino (2014). Short-term and long-term within-person variability in performance: An integrative model. *Journal of Applied Psychology*, 99(5), 898-914. doi: 10.1037/a0037402
- Minbashian, A., Wood, R. E., and N. Beckmann (2010). Task-contingent conscientiousness as a unit of personality at work. *Journal of Applied Psychology*, 95(5), 793-806. doi: 10.1037/a0020016
- Morris, A. S. et al. (2007). The role of the family context in the development of emotion regulation. *Social Development*, 16, 362-388.
- Munafò, M. R., Zettler, J. I., and T. G. Clark (2007). Personality and smoking status: A meta-analysis. *Nicotine and Tobacco Research*, 9(3), 405-413. doi: 10.1080/14622200701188851
- National Academy of Sciences (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, D.C.: National Academies Press.

- Neff, K. D., Kirkpatrick, K. L., and S. S. Rude (2007). Self-compassion and adaptive psychological functioning. *Journal of Research in Personality* 41 (2007) 139–154
- Nel, J. A. et al. (2012). Exploring the personality structure in the 11 languages of South Africa. *Journal of Personality*, 80(4), 915-948. doi: 10.1111/j.1467-6494.2011.00751.x
- Nigg, J. T. et al. (2002). Big Five dimensions and ADHD symptoms: Links between personality traits and clinical symptoms. *Journal of Personality and Social Psychology*, 83, 451-469.
- Obschonka, M. et al. (2013). The regional distribution and correlates of an entrepreneurship-prone personality profile in the United States, Germany, and the United Kingdom: A socioecological perspective. *Journal of Personality and Social Psychology*, 105(1), 104-122. doi: 10.1037/a0032275
- Obschonka, M., Schmitt-Rodermund, E., and A. Terracciano (2014). Personality and the gender gap in self-employment: A multi-nation study. *Plos One*, 9(8). doi: e103805
- Obschonka, M., Silbereisen, R. K., and E. Schmitt-Rodermund (2012). Explaining entrepreneurial behavior: Dispositional personality traits, growth of personal entrepreneurial resources, and business idea generation. *Career Development Quarterly*, 60(2), 178-190. doi: 10.1002/j.2161-0045.2012.00015.x
- OECD (2014). Skills for social progress. Paris, France: OECD-Centre for Education Research and Innovation.
- OECD (2015). Draft proposal: OECD Longitudinal Study of Social and Emotional skills in Cities. Paris: OECD.
- OECD (2015). ESP Longitudinal Study of Skill Dynamics in Cities. Paris: OECD.
- Oh, I. S., Wang, G., and M. K. Mount (2010). Validity of observer ratings of the Five-Factor Model of personality traits: A meta-analysis. *Journal of Applied Psychology*, 96(4), 762-773. doi: 10.1037/a0021832
- Omoto, A. M., Snyder, M., and J. D. Hackett (2010). Personality and motivational antecedents of activism and civic engagement. *Journal of Personality*, 78(6), 1703-1734. doi: 10.1111/j.1467-6494.2010.00667.x
- Oswald, A. J., Proto, E., and D. Sgroi (2014). Happiness and productivity. Warwick: University of Warwick.
- Papachristou, H. et al. (2013). Higher levels of trait impulsiveness and a less effective response inhibition are linked to more intense cue-elicited craving for alcohol in alcohol-dependent patients. *Psychopharmacology*, 228(4), 641-649. doi: 10.1007/s00213-013-3063-3
- Park-Higgerson, H. K. et al. (2008). The evaluation of school-based violence prevention programs: A meta-analysis. *Journal of School Health*, 78(9), 465-479. doi: 10.1111/j.1746-1561.2008.00332.x
- Poropat, A. E. (2014a). A meta-analysis of adult-rated child personality and academic performance in primary education. [Article]. *British Journal of Educational Psychology*, 84(2), 239-252. doi: 10.1111/bjep.12019

- Poropat, A. E. (2014b). Other-rated personality and academic performance: Evidence and implications. *Learning and Individual Differences*, 34 (2014), 24-32.
- Poropat, A. E. (2009). A Meta-Analysis of the Five-Factor Model of Personality and Academic Performance. *Psychological Bulletin*, 135(2), 322-338. doi: 10.1037/a0014996
- Primi, R. et al. (2014, submitted). The development of a nationwide inventory assessing social and emotional skills in Brazilian youth.
- Prinzie, P., Onghena, P., and W. Hellinckx (2005). Parent and child personality traits and children's externalizing problem behavior from age 4 to 9 years: A cohort-sequential latent growth curve analysis. *Merrill-Palmer Quarterly-Journal of Developmental Psychology*, 51(3), 335-366.
- Rescorla, L. A. et al. (2014). Parent-teacher agreement on children's problems in 21 societies. *Journal of Clinical Child and Adolescent Psychology*, 43(4), 627-642. doi: 10.1080/15374416.2014.900719
- Roberts, B. W., Walton, K. E., and W. Viechtbauer (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. doi: 10.1037/0033-2909.132.1.1
- Roberts, B. W., Wood, D., and J. L. Smith (2005). Evaluating Five Factor Theory and social investment perspectives on personality trait development. [Proceedings Paper]. *Journal of Research in Personality*, 39(1), 166-184. doi: 10.1016/j.jrp.2004.08.002
- Rosander, P., and M. Backstrom (2012). The unique contribution of learning approaches to academic performance, after controlling for IQ and personality: Are there gender differences? *Learning and Individual Differences*, 22(6), 820-826. doi: 10.1016/j.lindif.2012.05.011
- Saarni, C. (1999). *The development of emotional competence*. New York: Guilford Press.
- Saarni, C. (2011). Emotional development in childhood. In R. E. Tremblay, R. G. Barr, R. DeV. Peters, and M. Boivin (Eds.), *Encyclopedia on early childhood development: Emotions* (pp. 1-7). Montreal, Quebec: Centre of Excellence for Early Childhood Development and Strategic Knowledge Cluster on Early Child Development.
- Salgado, J. F. et al. (2003). International validity generalization of GMA and cognitive abilities: A European community meta-analysis. *Personnel Psychology*, 56(3), 573-605.
- Salgado, J. F. et al. (2003). A meta-analytic study of general mental ability validity for different occupations in the European community. *Journal of Applied Psychology*, 88(6), 1068-1081.
- Salovey, P. et al. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the trait meta-mood scale. In J. W. Pennebaker (Ed.), *Emotion, disclosure, and health* (pp. 125-154). Washington, DC: American Psychological Association.
- Santos, D. and R. Primi (2014). Promoting social and emotional skills for societal progress in Rio de Janeiro. Paris, France: OECD-Centre for Education Research and Innovation.
- Saucier, G., and F. Ostendorf (1999). Hierarchical subcomponents of the Big Five personality factors: A cross-language replication. *Journal of Personality and Social Psychology*, 76, 4, 613-627.

- Schmitt, D. P. et al. (2007). The geographic distribution of big five personality traits - Patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology*, 38(2), 173-212. doi: 10.1177/0022022106297299
- Schnittker, J. and J. R. Behrman (2012). Learning to do well or learning to do good? Estimating the effects of schooling on civic engagement, social cohesion, and labor market outcomes in the presence of endowments. *Social Science Research*, 41(2), 306-320. doi: 10.1016/j.ssresearch.2011.11.010
- Schraw, G., and R. S. Dennison (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19(4), 460-475. doi: 10.1006/ceps.1994.1033
- Scott, S., Briskman, J., and T. G. O'Connor (2014). Early prevention of antisocial personality: Long-term follow-up of two randomized controlled trials comparing indicated and selective approaches. *American Journal of Psychiatry*, 171(6), 649-657. doi: 10.1176/appi.ajp.2014.13050697
- Seligman, M. E. P. et al. (2005). Positive Psychology Progress: Empirical Validation of Interventions. *American Psychologist*, 60, 5, 410-421 DOI: 10.1037/0003-066X.60.5.410.
- Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin*, 20(5), 580-591. doi: 10.1177/0146167294205014
- Sklad, M. et al. (2012). Effectiveness of school-based universal social, emotional, and behavioral programs: do they enhance students' development in the area of skill, behavior, and adjustment? *Psychology in the Schools*, 49(9), 892-909. doi: 10.1002/pits.21641
- Soto, C. J. et al. (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.
- Soto, C. J. et al. (2008). The developmental psychometrics of Big Five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, 94, 718-737.
- Soto, C. J., and O. P. John (in press). Traits in transition: The structure of parent-reported personality traits from early childhood to early adulthood. *Journal of Personality*.
- Soto, C. J., and O. P. John (2015). *Conceptualization, development, and validation of the BFI-2: The next version of the Big Five Inventory*. Manuscript submitted for publication.
- Soto, C. J., and O. P. John (2012). Development of Big Five domains and facets in adulthood: Mean-level age trends and broadly versus narrowly acting mechanisms. *Journal of Personality*, 80, 881-914.
- Soto, C. J., and O. P. John (2009). Ten facet scales for the Big Five Inventory: Convergence with NEO PI-R facets, self-peer agreement, and discriminant validity. *Journal of Research in Personality*, 43, 84-90.
- Spengler, M. et al. (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
- Srivastava, S. et al. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology*, 84, 1041-1053.

- Stautz, K. and A. Cooper (2013). Impulsivity-related personality traits and adolescent alcohol use: A meta-analytic review. *Clinical Psychology Review*, 33(4), 574-592. doi: 10.1016/j.cpr.2013.03.003
- Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin*, 133(1), 65-94. doi: 10.1037/0033-2909.133.1.65
- Steel, P., Schmidt, J., and J. Shultz (2008). Refining the relationship between personality and subjective well-being. *Psychological Bulletin*, 134(1), 138-161. doi: 10.1037/0033-2909.134.1.138
- Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence*, 35(5), 401-426. doi: 10.1016/j.intell.2006.09.004
- Sutin, A. R. et al. (2011). Personality and obesity across the adult life span. *Journal of Personality and Social Psychology*, 101(3), 579-592. doi: 10.1037/a0024286
- Tackett, J. L. (2006). Evaluating models of the personality-psychopathology relationship in children and adolescents. *Clinical Psychology Review*, 26(5), 584-599.
- Tackett, J. L. et al. (2013). Delineating personality traits in childhood and adolescence: Associations across measures, temperament, and behavioral problems. *Assessment*, 20(6), 738-751. doi: 10.1177/1073191113509686
- Tackett, J. L. et al. (2013). Common genetic influences on negative emotionality and a general psychopathology factor in childhood and adolescence. *Journal of Abnormal Psychology*, 122(4), 1142-1153. doi: 10.1037/a0034151
- Tamir, M. et al. (2007). Implicit theories of emotion: Affective and social outcomes across a major life transition. *Journal of Personality and Social Psychology*, 92, 731-744.
- Taylor, N., and G. P. De Bruin (2005). *Basic Traits Inventory: Technical manual*. Johannesburg: Jopie van Rooyen and Partners, SA.
- Terracciano, A. et al. (2013). Personality and resilience to Alzheimer's disease neuropathology: a prospective autopsy study. [Article]. *Neurobiology of Aging*, 34(4), 1045-1050. doi: 10.1016/j.neurobiolaging.2012.08.008
- Thomas, A. et al. (1963). *Behavioral individuality in early childhood*. New York: New York University Press.
- Trapnell, P. D., and J. D. Campbell (1999). Private self-consciousness and the five-factor model of personality: Distinguishing rumination from reflection. *Journal of Personality and Social Psychology*, 76(2), 284-304. doi: 10.1037/0022-3514.76.2.284
- Trilling, B., and C. Fadel (2009). *21st century skills: Learning for life in our times*. San Francisco, CA: Jossey-Bass.
- Tsai, J. L., Knutson, B., and H. H. Fung (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology*, 90(2), 288-307. doi: 10.1037/0022-3514.90.2.288
- Valchev, V. H. et al. (2014). Beyond agreeableness: Social-relational personality concepts from an indigenous and cross-cultural perspective. *Journal of Research in Personality*, 48, 17-32. doi: 10.1016/j.jrp.2013.10.003

- Van de Vijver, F., and K. Leung (1997). *Methods and data analysis for cross-cultural research*. Thousands Oaks, CA: Sage.
- Van den Akker, A. L. et al. (2013). The Development of Personality Extremity From Childhood to Adolescence: Relations to Internalizing and Externalizing Problems. *Journal of Personality and Social Psychology*, 105(6), 1038-1048. doi: 10.1037/a0034441
- Van den Akker, A. L., Dekovic, M., and P. Prinzie (2010). Transitioning to adolescence: How changes in child personality and overreactive parenting predict adolescent adjustment problems. *Development and Psychopathology*, 22(1), 151-163. doi: 10.1017/s0954579409990320
- Van Hiel, A., Mervielde, I., and F. De Fruyt (2006). Stagnation and generativity: Structure, validity, and differential relationships with adaptive and maladaptive personality. *Journal of Personality*, 74(2), 543-573.
- Vandenberg, R. J., and C. E. Lance (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for roganizational research *Organizational Research Methods*(3), 4-70.
- Vecchione, M. et al. (2014). Fakability of Implicit and Explicit Measures of the Big Five: Research findings from organizational settings. *International Journal of Selection and Assessment*, 22(2), 211-218. doi: 10.1111/ijsa.12070
- Versteeg, H. et al. (2012). Type D personality and health status in cardiovascular disease populations: a meta-analysis of prospective studies. *European Journal of Preventive Cardiology*, 19(6), 1373-1380. doi: 10.1177/1741826711425338
- Wang, L., et al. (2009). Assessing Teamwork and Collaboration in High School Students: A Multimethod Approach. *Canadian Journal of School Psychology* 2009; 24; 108 DOI: 10.1177/0829573509335470.
- Weinberg, A., and E. Klonsky (2009). Measurement of emotion dysregulation in adolescents. *Psychological Assessment*, 21, 616–621.
- Wiggins, J. S. (1973). *Personality and prediction*. Jossey-Bass Publishers.
- Wille, B., De Fruyt, F., and M. Feys (2013). Big Five traits and intrinsic success in the new career era: A 15-year longitudinal study on employability and work-family conflict. *Applied Psychology-an International Review-Psychologie Appliquee-Revue Internationale*, 62(1), 124-156. doi: 10.1111/j.1464-0597.2012.00516.x
- Wille, B., De Fruyt, F., and M. Feys (2010). Vocational interests and Big Five traits as predictors of job instability. *Journal of Vocational Behavior*, 76(3), 547-558. doi: 10.1016/j.jvb.2010.01.007
- Woods, S. A. et al. (2013). Personality across working life: The longitudinal and reciprocal influences of personality on work. *Journal of Organizational Behavior*, 34, S7-S25. doi: 10.1002/job.1863