School engagement and intentional self-regulation
A reciprocal relation in adolescence

Kristján Ketill Stefánsson

UNIVERSITY OF ICELAND
SCHOOL OF EDUCATION

School engagement and intentional self-regulation

A reciprocal relation in adolescence

Kristján Ketill Stefánsson

Supervisor
Dr. Steinunn Gestsdóttir

Doctoral committee
Dr. Freyja Birgisdóttir
Dr. Richard M. Lerner
Dr. Sigurgrímur Skúlason

Opponents at defence
Dr. Álfgeir Logi Kristjánsson
Dr. Katariina Salmela-Aro

Dissertation submitted in partial fulfilment of a Ph.D.-degree

Faculty of Education Studies, School of Education and Faculty of Psychology,
School of Health Sciences, University of Iceland
May 2017

School engagement and intentional self-regulation
A reciprocal relation in adolescence

A thesis for a Ph.D.-degree in Educational sciences

Preface

The current Ph.D. dissertation (180 ECTS) is a collection of three scientific articles and a synopsis combining the articles into a comprehensive scientific entity. I have already published two of these articles. The first article "The development and validation of the SOC scale for youth in Iceland" was published in 2014 in Sálfræðiritið (Journal of the Icelandic Psychological Association). The second article "Bifactor Model of School Engagement: Assessing General and Specific Aspects of Behavioral, Emotional and Cognitive Engagement among Adolescents" was published in 2016 in the International Journal of Behavioral Development. A manuscript of the third article "School engagement and intentional self-regulation: A reciprocal relation in adolescence" was submitted to an international refereed journal before the dissertation defence. This dissertation is the result of my own work and includes nothing that is the outcome of work done in collaboration except as specified in the text. My main-supervisor during the docotral studies was Dr. Steinunn Gestsdóttir, and the co-supervisor was Dr. Freyja Birgisdóttir. Dr. Richard M. Lerner and Dr. Sigurgrímur Skúlason joined the doctoral committee at later stages.

Without financial support, I would not have managed to finish my studies. I would like to acknowledge the financial support obtained by my supervisor from the University of Iceland Research Fund and University of Iceland Doctoral Fund that made it possible for me to devote three concentrated years of work to this study. I owe gratitude to many people for helping me develop the research skills that I now have. Some of these influential people I have not seen or worked with for years, but still they play a significant part in the work described in this synopsis. I will not try to name all of these people but I will try to mention the most significant groups that come to mind. First, I want to thank all the students I have ever taught. Second, I want to thank my colleagues at Laugalækjarskóli secondary school, the School of Education, and the Department of Psychology. Third, I want to thank the study participants, my doctoral committee, my opponents, and my supervisor. Finally, and most importantly, I want to thank my family and friends, without you my models would never converge.

Reykjavík, May 2017, Kristján Ketill Stefánsson

Abstract

The importance of school engagement (i.e., the willingness to engage in learning) for school success, such as good academic achievement and low dropout rates, has been well established. At the same time, intentional self-regulation (ISR; i.e., the ability to set, prioritize, and obtain long-term goals) has been shown to be a precursor, mediator, and outcome of school engagement. However, the relation between school engagement and ISR during adolescence is poorly understood. In this research, I explored the reciprocal relation between school engagement and ISR during adolescence. This study had three goals. The first goal was to further the development of a valid measure of ISR for use with adolescents. The second goal was to contribute to the development of a valid measure of school engagement for use with adolescents. The third goal, which best captures the main purpose of the study, was to examine the hypothesized reciprocal relation of school engagement and ISR during the last two years of compulsory school in Iceland. The results from the development and adaptation of the school engagement and ISR measures were published in two journal articles based on four waves of data collected at the beginning and end of Grades 9 and 10 with a longitudinal sample of 561 youth in Iceland (46% girls, M_{age} at Wave 1 = 14.3 years, SD = 0.3). The third and final manuscript, based on data from the same longitudinal sample, supported the reciprocal relations of school engagement and ISR during adolescence after controlling for gender, academic achievement, and parent's education. Furthermore, the results indicated decreased stability of both school engagement and ISR during the observed period. The decreasing stability is consistent with theories that present school engagement and ISR as malleable constructs that are open to contextual conditions. The reciprocal relations between school engagement and ISR support hypotheses that ISR skills are a key element in the promotion of school engagement.

Keywords: *intentional self-regulation, school engagement, positive youth development, adolescence, SOC.*

Abstract in Icelandic

Virk þátttaka í skólastarfi og sjálfstjórnun: Gagnvirkt samband á unglingsárum

Virk þátttaka í skólastarfi (e. *school engagement; skuldbinding til náms*) er mikilvæg fyrir farsæla skólagöngu en sýnt hefur verið fram á skýr tengsl virkrar þátttöku í skólastarfi og jákvæðra þátta eins og t.d. hárra einkunna og lítillar hættu á brottfalli. Á sama tíma hefur verið sýnt fram á að sjálfstjórnun (e. *intentional self regulation; hæfileikinn til að setja sér, forgangsraða og ná langtímamarkmiðum*) spáir fyrir, miðlar og verður fyrir áhrifum af virkri þátttöku nemenda. Þrátt fyrir það er lítið vitað í hvaða röð þessi tengsl eiga sér stað á unglingsárum. Í þessari rannsókn kannaði ég möguleikann á hvort að jákvætt gagnvirkt samband gæti verið til staðar á milli virkrar þátttöku nemenda og sjálfstjórnunar á unglingsárum. Markmiðum rannsóknarinnar var skipt í þrennt. Fyrsta markmiðið var að auka við fyrirliggjandi þekkingu á því hvernig mæla má sjálfstjórnun á unglingsárum. Annað markmiðið var að auka við fyrirliggjandi þekkingu á því hvernig mæla má virka þátttöku í skólastarfi á unglingsárum. Þriðja markmiðið var að prófa tilgátuna um jákvætt gagnvirkt samband milli virkrar þátttöku nemenda og sjálfstjórnunar síðustu tvö ár grunnskólans. Niðurstöður úr aðlögun og þróun mælitækjanna á virkri þátttöku nemenda og sjálfstjórnun voru birtar í tveimur tímaritsgreinum sem byggðu á langtímagögnum sem safnað var frá 561 unglingi (46% stelpur, M_{aldur} við byrjun 9. bekkjar = 14,3 ár, *Staðalfrávik* = 0,3) við upphaf og lok 9. og 10. bekkjar. Niðurstöður úr lokagrein rannsóknarinnar byggðu á sömu langtímagögnum og studdu megintilgátuna um jákvætt gagnvirkt langtímasamband milli virkrar þátttöku og sjálfstjórnunar eftir að stjórnað hafði verið fyrir áhrifum kyns, fyrri námsárangurs og menntunar foreldra. Að auki sýndu niðurstöðurnar minnkandi stöðugleika virkar þátttöku og sjálfstjórnunar eftir því sem leið að lokum grunnskólans. Minnkandi stöðugleiki á tímabilinu er í samræmi við kenningar sem lýsa virkri þátttöku nemenda og sjálfstjórnun sem mótanlegum þáttum. Sambandið milli virkrar þátttöku nemenda og sjálfstjórnunar styður við tilgátur sem sýna sjálfstjórnun sem lykilhugtak í stuðningi við virka þátttöku í skólastarfi.

Table of Contents

List of figures

List of tables

List of papers

This theses is based upon the following papers, referred to by their Roman numerals

1   Review of the literature

The importance of school engagement (i.e., the willingness to engage in learning) for school success, such as good academic achievement and low dropout rates, has been well established (see e.g., Christenson, Reschly, & Wylie, 2012; Tuominen-Soini & Salmela-Aro, 2014). At the same time, recent studies have shown that half of American students are not engaged in school (e.g., Gallup Student Poll, 2015), which correspond to findings from cultures across the world (OECD, 2012a). Importantly, school engagement has been considered to be malleable and, as such, open to contextual conditions, including influences from parents, teachers, as well as students themselves (Appleton, Christenson, & Furlong, 2008; Fredricks, Blumenfeld, & Paris, 2004). Accordingly, researchers and educators have called for research that seeks to identify what promotes school engagement in the classroom (Coalition for Psychology in Schools and Education, 2006; Shernoff, 2013).

Intentional self-regulation (ISR) has been conceptualized as the ability to set, prioritize, and obtain long-term goals (Freund & Baltes, 2002) and has been suggested to be an important precursor, mediator, and outcome of school engagement (Cleary & Zimmerman, 2012; Wolters & Taylor, 2012). Scholars have called for a better understanding on the possible overlap between school engagement and ISR (Boekaerts, 2016; Eccles, 2016) and the hypothesized reciprocal relation between school engagement and ISR (Pintrich, 2003; Wolters, 2003). A better understanding is needed on how school engagement and ISR may work in tandem during adolescence, allowing positive effects of both constructs to cumulate and thereby propelling the student on a positive academic trajectory. The current study takes a step towards such an understanding by assessing a reciprocal relation between school engagement and ISR among adolescents in Iceland during the last two years of compulsory schooling.

Children in Iceland normally start compulsory education the year they turn six years old and progress automatically from one grade level to the next for 10 years until the year they turn 16 years old. After compulsory school, most students proceed to upper secondary school, although it is not compulsory. Upper secondary schools are three to four year programs that fall into three main categories: grammar schools, comprehensive schools, and vocational schools. The dropout rate at the upper secondary level in

Iceland is high. As an example, less than half (45%) of the students that registered in Icelandic upper secondary schools in the academic year 2002-2003 finished within the expected four years. After upper secondary school (around the age of 19-20 years) students can choose some form of university education. The university level is, in terms of the European Bologna framework for higher education, mostly a three year bachelor, a two year master, and a three year doctoral cycle system (Blondal, Jónasson, & Tannhäuser, 2011; OECD, 2012a). At this writing, the part of adolescence (Grades 9 through 10) examined in the current study is of special relevance to discussion of school dropout and the promotion of learning and achievement in Iceland. Grades 9 through 10 mark the last two years of compulsory school. After Grade 10, Icelandic students can choose which upper secondary school they would like to attend (comprehensive, grammar, or vocational), and admission in the most popular schools is commonly decided on the basis of grades from their last year of school only. Therefore, Grade 10 is generally viewed as a period of greater expectations and increased urgency with regard to educational goals.

There are three main theoretical perspectives used in this study. The importance placed on ISR for adolescent development stems from the relational developmental systems perspective (Overton, 2010, 2013, 2015). Furthermore, the current study is informed by the work of Lerner, Lerner and colleagues on positive youth development (PYD), where school engagement and ISR are regarded as key adolescent strengths that, together with ecological assets, promote the positive development of youth (Figure 1; Lerner, Lerner, & Benson, 2011). Finally, this study highlights the hypothesized reciprocal relation between school engagement and ISR as proposed by the theoretical model of motivational dynamics (Figure 2; Skinner & Pitzer, 2012) .

In the following sections, I start by describing the initial theoretical context that the current study is based on. Next, the constructs of ISR and school engagement are described, together with pressing issues regarding the operationalization of both constructs, which are addressed in Papers I and II. Next, I argue for the main hypothesis of the study, the reciprocal relation between school engagement and ISR, addressed in Paper III. Finally, the last section of this literature review describes the aims and significance of the study.

1.1 The origin of this dissertation: Why study adolescent strengths?

The idea for the current study can be traced to my personal experience as a science and mathematics teacher in Grades 9 and 10. In my work as a teacher, I came across attitudes of teachers, parents, and students that I experienced as dualistic in nature. These attitudes are best described by simple phrases like "will is all you need" or "the students just need to learn how to learn". Consistent with many developmental theories (e.g., Bandura, 1986), I thought I witnessed the importance of both the students' willingness to engage in learning and the skills they needed to achieve their academic goals. When I started my doctoral studies with Dr. Steinunn Gestsdottir, and joined her study on the positive youth development, I was able to put my personal thoughts about adolescent strengths, their engagement in school, and their well-being, into a theoretical context. As I progressed in my program, I realized that limited empirical research seemed to exist on the relation between motivation and self-regulation and the goal of my dissertation became to conduct a study where a testable hypothesis about this bidirectional relation could be made. In this chapter, I describe the theoretical perspectives that shaped my thinking about adolescent strengths that guided the formulation of the current study.

1.1.1 Relational developmental systems

The focus on the active, bidirectional interaction between ISR and school engagement in this study is partly informed by the process relational paradigm (Overton, 2015). According to this paradigm, adolescents (as well as other living organisms) are inherently active, self-creating, self-organizing, self-regulating, plastic, and non-linear complex adaptive systems. What this means, is that the development of adolescents is shaped by their own embodied activities and actions, which operate coactively in a lived world of physical and sociocultural objects (Overton, 2015).

When studying human development, a relational developmental system metatheory emphasizes the mutual bidirectional relations between the individual and the context in which he or she lives, as both progress over time (Brandtstädter, 1998, 2006). These relations determine the development of the individual and have the potential for systematic change. This potential for change is important for the study of adolescence, as it allows for the possibility that the individual and the context can be altered to increase the probability of a positive change in the development

of each individual (Lerner, Lerner, & Benson, 2011). As such, the mutual bidirectional relations between the individual and the context are important for the current study, as it highlights the potential of ISR to promote adaptive relations between the young person and his or her environment. If adaptive developmental relations between the developing adolescent and features of his or her school can be fostered, it increases the likelihood that the young person will thrive (Lerner, Lerner, Bowers, & Geldhof, 2015).

The systematic relations that adolescents have with key people and institutions while undergoing self-changes create the major source of diversity in the developmental trajectories of youth. These relations have different timings for different individuals and stem from biological, psychological, and societal factors, with no one factor acting in isolation (Lerner, Phelps, Forman, & Bowers, 2009). Acknowledging that systematic relations have different timings for different individuals is important for the current research as it guides the interpretation of the findings of the study away from making strict normative assumptions about the design of learning environments and encourages researchers and practitioner to design flexible learning environments that can accommodate people on different developmental trajectories (see Rose, 2015).

The diversity of individual developmental trajectories is of importance for the analysis of longitudinal data. The standard approach to longitudinal analyses has been to analyze data at the between subjects level and the implicit assumption has been that the results are applicable at the within subject-level. This assumption, known as the assumption of ergodicity, does not hold for many psychological processes (Molenaar, 2004). However, the fact that developmental processes are non-ergodic does not mean that between- subjects analysis is without meri; the extent to which group-level observations reflect person-level phenomena remains a still largely unexplored empirical issue (McClelland, Geldhof, Cameron, & Wanless, 2015). As the current research, like the majority of previous research on psychological processes, are limited to between subjects analysis, it has to be kept in mind that the results of the current study only apply reliably to the group-level.

1.1.2 Positive youth development

At the outset of this study I used the positive youth development (PYD) perspective to frame a discussion about the role of adolescent strengths (i.e., school engagement and ISR) for the positive development of youth. By

viewing school engagement, ISR, and learning environments as potential resources for teachers and students, attention is given to strengths and ecological assets that can be built upon to promote the positive development of youth. As such, the PYD view of development is helpful to identify youth strengths and provide a constructive starting point for the design of effective learning environments.

Since the early 1990s, several research traditions have focused on replacing the deficit view in research on adolescence with a more positive view of human development. The focus on positive outcomes is found both in positive psychology and in the study of resilience (Rutter, 2006; Seligman & Csikszentmihalyi, 2000). However, this body of work is independent of PYD, which has its origins in developmental science and comparative psychology (Lerner et al., 2013; Lerner et al., 2009; Lerner et al., 2015).
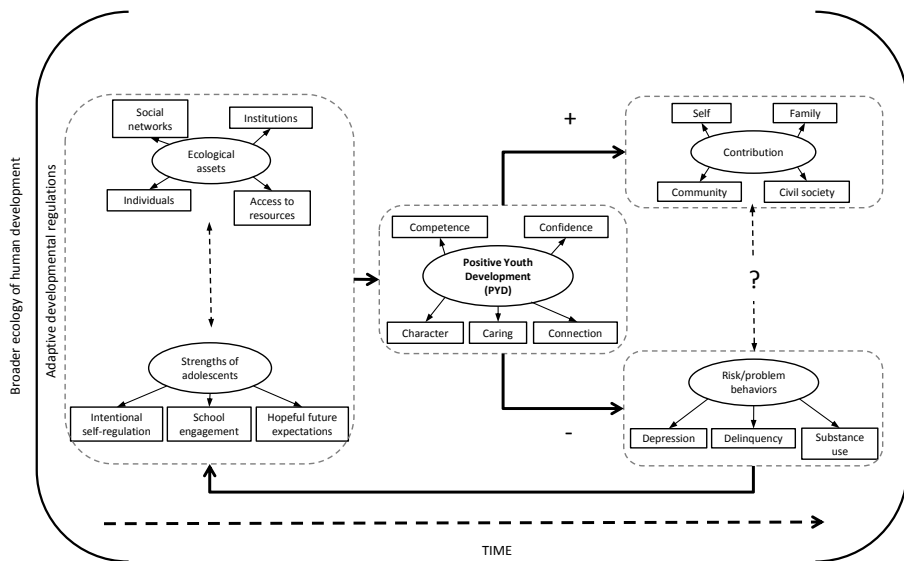
Figure 1. The relational developmental systems model of the individual ⟷ context relation involved in PYD (from Lerner, Lerner, & Benson, 2011).

The PYD perspective represents an orientation toward youth that emphasizes positive outcomes and the importance of individual plasticity, as well as the relations between individuals and their contexts as the basis of variation in the course of human development. As such, the PYD perspective is a dynamic model of human behavior and consistent with the relational developmental systems metatheory described in the previous section. The mutually inferential individual ⟷ context relations of the PYD

process, as depicted by Lerner, Lerner and colleagues, are shown in Figure 1. The figure shows how strengths of adolescents, when aligned with ecological assets, promote positive youth development, which in turn affects both the adaptive and problematic behaviors of adolescents (Lerner et al., 2009; Lerner, Lerner, & Benson, 2011).

Lerner and Lerner and colleagues proposed constructs of PYD, labeled the Five Cs (Lerner et al., 2005). The Five Cs are situated in the middle of Figure 1, and stem from the experience of practitioners and from reviews of the adolescent development literature (Eccles & Gootman, 2002; Roth & Brooks-Gunn, 2003). The Five Cs stand for: Competence, Confidence, Connection, Character, and Caring. Table 1 includes a brief definition of each of the Five Cs.

Table 1. A brief definition of the Five Cs of Positive Youth Development (from Lerner et al., 2009)

Competence: Positive view of one's actions in specific areas, including social, academic, cognitive, health, and vocational. Social competence refers to interpersonal skills (e.g., conflict resolution). Academic competence refers to school performance as shown, in part, by school grades, attendance, and test scores. Cognitive competence refers to cognitive abilities (e.g., decision making). Health competence involves using nutrition, exercise, and rest to keep oneself fit. Vocational competence involves work habits and explorations of career choices.

Confidence: An internal sense of overall positive self-worth and self-efficacy.

Connection: Positive bonds with people and institutions that are reflected in exchanges between the individual and his or her peers, family, school, and community in which both parties contribute to the relationship.

Character: Respect for societal and cultural norms, possession of standards for correct behaviors, a sense of right and wrong (morality), and integrity.

Caring/Compassion: A sense of sympathy and empathy for others.

The current study is focused on the relation between two individual strengths in adolescence, school engagement and ISR, both of which have been considered to be important antecedents of the Five Cs, which includes academic competence (see Table 1; Lerner, Lerner, & Benson, 2011). According to Lerner and Lerner's Five Cs model of PYD, school engagement

and ISR are important strengths in adolescence, which, when aligned with ecological assets, promote PYD (see Figure 1). Accordingly, knowledge about the hypothesized reciprocal relation between school engagement and ISR can help us understand how to support the healthy development of youth.

1.1.3 Empirical support for the study of adolescents' strengths and PYD

A major source of research that has provided empirical support for the PYD perspective comes from the 4-H Study of Positive Youth Development (Lerner et al., 2015). The 4-H study aimed at identifying the individual and ecological bases of healthy development among adolescents, as well as providing evidence for indicators of PYD. Richard M. Lerner and Jacqueline V. Lerner directed the 4-H study, which was launched in 2001. The study used a form of longitudinal sequential design in which 1700 fifth graders, during the 2002–2003 school year, were the initial cohort. After the first wave of measurement, the data set included information from 13 U.S. states. Data collection in the 4-H Study of Positive Youth Development was completed after the assessment of 12th grade youth in the 2010-2011 school year (for full details see for example, Lerner, 2005; Lerner, Lerner, von Eye, Bowers, & Lewin-Bizan, 2011; Lerner, von Eye, Lerner, Lewin-Bizan, & Bowers, 2010).

The researchers in the 4-H Study used a student questionnaire to collect information about PYD. Using data from Wave 1 (Grade 5) of the 4-H Study, structural equation modeling provided evidence for five first order latent factors representing the Five Cs of PYD and for their convergence on a second order PYD latent construct (Lerner et al., 2005). Furthermore, consequent studies have confirmed that PYD predicted higher youth contribution, lower risk behaviors, and depression at later grades. In addition, results have shown that promoting PYD is not equivalent to preventing risk/problem behaviors. Instead, a multiplicity of patterns and conjoint trajectories emerged for PYD and risk/problem behavior, creating a new perspective for conceptualizing PYD (Lerner et al., 2009; Zimmerman, Phelps, & Lerner, 2008).

The mutually influential individual ↔ context relations in the PYD process highlight the importance of considering whether PYD is comparable across different cultural contexts. The cultural relevance of PYD in Western societies, outside the U.S., has been discussed and partially supported in studies that have addressed theoretical issues of PYD (Silbereisen & Lerner,

2007), the measurement of ISR (Gestsdottir et al., 2015), and the measurement of the Five Cs (Holsen, Geldhof, Larsen, & Aardal, in press). The cultural relevance of PYD in societies outside the U.S. is an important argument for applying the PYD perspective in Iceland.

Researchers have recommended several steps for future research regarding PYD. For example, although this is a growing research area (see e.g., Wen, Su, Li, & Lin, 2015), little is known about the cultural relevance of PYD outside Western societies, and more research is needed on the factor structure and mean levels of PYD in different groups, such as among boys and girls. Furthermore, researchers have recommended that the scope of the PYD context should be broadened to include the context of adolescents, such as schools, where young people spend a large proportion of their time (Lerner et al., 2009). These recommendations are important for the current study, as they are arguments for studying antecedents of the Five Cs in the context of school in Iceland.

1.2 Paper I: Defining and measuring intentional self-regulation (ISR)

The importance placed on ISR in the mutual individual ↔ context relation involved in PYD, described earlier, guided my research efforts at the beginning of this study and led to the writing of the first journal article (Paper I). This paper was aimed at clarifying conceptual and methodological issues concerning the measurement of ISR. The following section describes the theoretical and empirical work that directed the study described in Paper I.

1.2.1 Intentional self-regulation (ISR)

The previously reviewed relational developmental systems metatheory and the 5 Cs of PYD model (see Figure 1) have highlighted the active role of the individual in his or her own developmental trajectory and identified ISR as one of the key strengths that help adolescents promote their own positive development. ISR has been defined as goal-directed behaviors aimed at harmonizing demands and resources in the environment with personal goals (Gestsdottir & Lerner, 2008). ISR allows people to set, prioritize, and obtain long-term goals and promote self-development (Freund & Baltes, 2002). Self-regulation can be separated into organismic and intentional self-regulation. Processes of organismic regulation are biologically based structures that are under little or no control of the person involved (e. g. circadian rhythms and pubertal timing). Intentional self-regulation (ISR) are

processes that are more readily available to consciousness and are amendable to control by the individual (Gestsdottir & Lerner, 2008).

Evidence from different fields suggests that biological, cognitive, and contextual changes in adolescence promote the development of ISR, that ISR becomes more focused and complex during this time, and that ISR can be used to achieve long-term goals in adolescence and adulthood (Brandtstädter, 2006). As such, ISR has been suggested as crucial component to healthy development during adolescence (Gestsdottir & Lerner, 2008). These suggestions are important for the current study, as they highlight the importance of studying ISR during a period marked by increased expectations of student self-sufficiency and increased social expectations in commitment to educational goals (Heckhausen & Tomasik, 2002).

Multiple examples of specific goal-directed behaviors, often directed at short-term goals, can be found in the literature, for example: learning strategies (Zimmerman, 2002), strategies to maintain motivation in education (Wolters, 2003), adaptive help-seeking (Newman, 2002), or other knowledge-based pragmatics, such as practicing mnemonic techniques to become a memory expert (Baltes & Baltes, 1990). The current study assessed ISR as goal-directed behaviors aimed at long-term goals that can be applied throughout the life span. The following section describes the rationale for using, and the general framework for understanding, ISR in the current study.

1.2.2 The SOC model

In the current research, the Selection, Optimization, and Compensation (SOC) model (Baltes, 1997; Baltes & Baltes, 1990) was used as a general framework for understanding and measuring ISR. The main reason for choosing the SOC model over other ISR conceptions is that SOC can be applied across the life span and is therefore suitable for longitudinal comparison. Furthermore, the SOC model has been used in the previously described 4-H Study in the U.S. (e.g., Lerner et al., 2005). The access to previous research results was important for the current study and made the operationalization of ISR in Iceland, using the SOC model feasible. In the following paragraphs, I provide an overview about the theory and measurement of SOC.

According to the SOC model, SOC consists of the orchestration of three component processes, selection (S), optimization (O), and compensation (C), plays an important role in acquiring developmentally relevant

resources. More specifically, selection consists of two sub-processes, elective selection (ES), and loss-based selection (LBS; see Table 2).

Table 2. Selection, optimization, and compensation embedded in an action-theoretical framework (Freund & Baltes, 1998)

Selection (goals / preferences)	Optimization (goal relevant means)	Compensation (means / resources for counteracting loss / decline in goal relevant means)
elective selection • specification of goals • goal system (hierarchy) • contextualization of goals • goal-commitment loss-based selection • focusing on most important goals • reconstruction of goal hierarchy • adaptation of standards • search for new goals	• attentional focus • seizing the right moment • persistence • acquiring new skills and resources • practice of skills • effort/energy • time allocation • modeling successful others	• substitution of means • use of external aids/help of others • use of therapeutic intervention • acquiring new skills/resources • activation of unused skills/resources • increased effort/energy • increased time allocation • modeling successful others who compensate • neglect of optimizing other means

The SOC principles are important for the current study as they form a theoretical base to discuss what characterizes ISR during the period of adolescence. The first sub-process, elective selection, deals with the specification, contextualization, commitment and hierarchy of goals. An example of high elective selection ability would involve an adolescent who has a clear hierarchy of what he or she wants to achieve in life, knows what

goals he or she wants and in what order, and is committed to the path chosen.

The second sub-process of selection is loss-based selection. Loss-based selection is a SOC strategy that is closely related to compensatory actions, the difference being in the adjustment of goals rather than means (i.e., compensation) when faced with obstacles in goal achievement. An example of high loss-based selection ability, relevant to the current study, is an adolescent who is able to select an alternative course of study when faced with less than adequate grades in a subject where grades were previously high.

The optimization (O) process is the use of internal and external resources as means to achieve goals previously selected. An example of an adolescent with high optimization ability would be an adolescent who is able to allocate time for homework and devote effort and energy to finishing his/her homework on time. Previous research has shown the optimization component process to be a strong manifestation of general ISR in adolescence (Bowers, Wang, Tirrell, & Lerner, 2016).

Finally, compensation (C) refers to the substitution of goal relevant means when the means are no longer available due to lack of resources. An example of an adolescent with high compensation ability, relevant to the current study, would be an adolescent who is able to acquire new learning strategies, such as summarization skills, note taking, or test anticipation skills, when prior strategies prove to be insufficient to reach a selected academic goal (Baltes, 1997; Baltes & Baltes, 1990).

According to the SOC model, the realization of the SOC components is dependent on the specific personal and societal circumstances of the individual as they get older (Baltes & Baltes, 1990). In other words, different personal and/or societal circumstances affect the way individuals perceive and use selection, optimization, and compensation for successful development. This point is important for the current research, as it highlights the need to look closely at the development of ISR during different periods of life, such as adolescence, when personal and societal circumstances undergo significant changes.

1.2.3 Measuring SOC

ISR in this study was measured using the selection, optimization, and compensation measure (SOC). However, the measurement of SOC has several unresolved issues regarding dimensionality, reliability, and validity in adolescence. Therefore, the development and assessment of a reliable

SOC measure for adolescents in Iceland was a major part of the current study (Paper I). Accordingly, I describe the measurement issues of SOC and the development of the SOC measure in some detail.

The life management strategy of selection, optimization, and compensation was originally operationalized in a study using a self-report measure (Baltes, Baltes, Freund, & Lang, 1999). In the original version of the SOC measure, each SOC construct was measured by 12 items, each using a two-statement forced-choice format. Thus, the questionnaire (including the loss-based selection process) consisted of 48 statements indicating a SOC related behavior and 48 statements indicating non-SOC related behavior. The expected four-factor solution was later confirmed in two independent adults samples in Germany (N = 218, 14-87 years; N = 181, 18-89 years). The studies showed convincing convergent and divergent correlations with other psychological constructs and showed moderate positive correlations with indicators of successful life management (Freund & Baltes, 2002).

The 4-H Study of PYD described previously included research on the SOC skills of adolescents (Lerner et al., 2005; Lerner, Lerner, & Benson, 2011). The SOC measure in the 4-H study during the first two waves (Grades 5 and 6) consisted of a short, six-item per subscale version. During the first two waves of the study, the loss-based selection scale (LBS) was not included, as LBS was not considered relevant during a period of substantial growth. The measurement tool therefore consisted of 18 items (three scales, six items per scale) that had shown adequate psychometric properties (Baltes et al., 1999).

The results from the first two waves of the 4-H Study did not reveal the three-factor structure of the measurement tool found in previous research with older participants in Germany. Instead, the researchers confirmed a nine-item global structure of SOC containing two selection items, four optimization items, and three compensation items (see Table 10; Gestsdottir & Lerner, 2007). The internal consistency of the nine-item structure was greater during the second wave (α = .64) than the first (α = .55). The nine-item measure had a maximum score of nine and a minimum score of zero. At Wave 1, the overall sample had an average SOC score of 6.63 (N = 1619) and a standard deviation of 1.88. At Wave 2, the average score was 6.51, (N = 1563) and the standard deviation was 2.02. The global structure showed weak to moderate correlations to indicators of positive and negative development in the expected directions (Gestsdottir & Lerner, 2007). The information about the lack of a three-factor structure in Grades 5 and 6 and the moderate correlation of a global factor to indicators of

positive and negative development is important for the current study, as it provided a reason to examine the SOC measurement carefully before it was used with adolescents in Iceland.

The three-factor structure of SOC was supported in a second longitudinal study with young people in Grades 8, 9, and 10 (N = 937) from the 4-H Study. The results also showed that the differentiation between the SOC strategies became more distinct across grades. The global SOC strategies showed a considerable linear decline, dropping close to a half standard deviation over the five-year period that had passed since the first measurement. As compared to optimization and compensation, elective selection in the eighth and ninth grades did not predict PYD in the tenth grade. The nine-item global ISR, on the other hand, positively predicted PYD and negatively predicted indicators of negative development to a moderate degree (Gestsdottir, Lewin-Bizan, von Eye, Lerner, & Lerner, 2009). The emergence of a three-factor structure in Grades 8, 9, and 10 was important for the present study, as it drew attention to the possibility that the SOC measurement might be developed further to create a more developmentally-sensitive measure for use with adolescents in Iceland. Accordingly, the adaptation of an Icelandic ISR measure, based on the SOC model, was another important goal of the current study.

As previously mentioned, the LBS subscale was not used during the first waves of data collection in the 4-H Study, as it was not considered developmentally relevant during an age period primarily focused on growth. After the sixth wave of data collection, researchers in the 4-H Study confirmed a four-part structure of ISR, including the LBS scale. The optimization factor was the only part of the SOC strategies to show a strong relationship (r = .43) to indicators of positive or negative development. The strong relationship was with the composite index of PYD (Gestsdottir, Bowers, von Eye, Napolitano, & Lerner, 2010). These results had relevance to the current study, as they showed that LBS has formed a subscale with adolescents in the U.S. and might therefore do so as well with Icelandic adolescents.

Consistent with the relational development metatheory, Zimmerman et al. (2008) investigated the relationship between SOC strategies and indicators of positive and negative development using a more person-centered approach. Five PYD trajectories represented change across Grades 5 through 8; for example, a trajectory that showed a linear rise in PYD through Grades 5 to 8 consisted of 21.3% of the sample, whereas the rest of the sample indicated a decline in PYD scores, to various levels, throughout

the age period. Similarly, 13.3% of the sample showed a considerable linear rise in risk behaviors, whereas the large majority did not show a considerable rise in risk behaviors. Binominal logistic regression indicated that young people with high SOC scores were somewhat more likely to be in the optimal trajectories (Zimmerman et al., 2008).

Similarly, using growth mixture modeling with data from the first seven waves of the 4-H Study, Bowers et al. (2011) identified four group trajectories of SOC. The largest group consisted of students having steadily declining SOC scores during the period (82%). The other three groups were composed of students experiencing elevated (8%), late onset (5%), and pronounced (5%) decline. The authors reported that the normative development in the sample was a general decline in self-reported ISR over the course of adolescence. The authors furthermore highlighted that the pattern of change was similar to reported declines in other psychological attributes in adolescence, such as grades, intrinsic motivation, self-concepts, and self-perceptions, as well as confidence in one's intellectual abilities (Bowers et al., 2011).

As previous studies have not identified a well-established and reliable tripartite structure of SOC in adolescence, researchers involved in the 4-H Study sought to find means to identify better ways to apply the SOC model and measure within the study of adolescence. Exploratory factor analysis using data from the eight waves of data from adolescents in the 4-H Study revealed, in addition to a general SOC factor, a reverse-coded method factor. This finding suggests that the format of some of the SOC items may be problematic, for American youth at least, and that future research should consider modifications to the measure to create a more developmentally sensitive measure for use with adolescents (Geldhof, Bowers, Gestsdottir, Napolitano, & Lerner, 2015).

Scale development research with U.S. data using the SOC questionnaire has furthermore recommended presenting the SOC items using a Likert-type scale, instead of using a forced-choice format, as a means to increase measurement precision (Geldhof, Gestsdottir, et al., 2015; Geldhof, Little, & Hawley, 2012). A part of the scale development research with U.S. data was co-authored by the doctoral candidate during the doctoral studies (see Geldhof, Gestsdottir, et al., 2015).

1.2.4   SOC research with adolescents in Iceland

The forced-choice SOC measure has been used with Icelandic adolescents in one prior study (Gestsdottir, Adalbjarnardottir, & Thorsdottir, 2011). The

data collection was conducted in the year 2009 with 505 students born in 1995 (9th graders) and 533 students born in 1991 (18-year-old students in upper secondary school). An 18-item version of the SOC measure was used and involved six items per component (Baltes et al., 1999). The researchers did not find the three-factor structure of SOC in the two samples. However, the researchers confirmed a seven-item global structure among the 14-year-olds and a nine-item global structure among the 18-year-olds. All of the items in the seven-item structure were also in the nine-item structure. Furthermore, all the items in the nine-item global structure corresponded to the nine-item global structure that had been confirmed previously with data from adolescents in the United States (see Table 10; Gestsdottir et al., 2011; Gestsdottir & Lerner, 2007; Zimmerman, Phelps, & Lerner, 2007). The results indicated that the subscales in the 18-item measure of SOC did not exist for a sample of 14- and 18-year-olds. The lack of a tripartite structure among late adolescents, in particular, called for a further assessment and development of the measurement with Icelandic samples.

1.2.5 Assessing the validity of SOC

The assessment and development of the SOC measure in the current study called for a related ISR measure to gauge the validity of SOC. I chose a measure called self-efficacy for self-regulated learning (SRL). The SRL measure, like the SOC measure, is intended to capture ISR. SRL is an aptitude measure of the self-directive processes and self-beliefs that enable learners to transform their mental abilities into academic performance (see Appendix A; Bandura, 2006; Zimmerman, 2008). The SRL measure is distinct from the SOC measure, as the SRL measure is directed at specific means towards learning goals, whereas SOC is directed at the optimization of means in general, and the management of goals and means in general (see Appendix B).

SOC strategies and SRL abilities are related processes with theoretically distinct features. This disparity is useful for the current study, as a related measure was needed to validate the measure of SOC among adolescents. I chose SOC rather than SRL as the main measure of ISR in the current study. There were three reasons for this. First, the SOC measure is more general and has more relevance for different goal-directed behaviors. Second, the SOC measure has a life-span perspective that gives the measure a greater temporal range and therefore applicability across different periods and contexts of life. Third, the access to previous research findings using the SOC measure created an opportunity to validate some of our findings.

Investigating the factor structure and the convergent validity of a multidimensional measure of ISR skills (i.e., the SOC measure) was the main aim of Paper I.

1.3 Paper II: Defining and measuring school engagement

The second paper was aimed at clarifying theoretical and methodological issues regarding the second main construct under investigation in this doctoral project, school engagement. This section describes the theoretical and empirical literature that laid the foundation for Paper II.

1.3.1 Defining school engagement

Research on school engagement[1] gained a momentum at the start of the new millennium. In fact, one of the biggest databases on psychological research (Psychinfo), recorded more than 32,000 articles on engagement from 2001-2015 (Azevedo, 2015). However, the research field still suffers from conceptual confusion, as different researchers label different phenomena as "engagement" or use different labels to describe the same concept (Reschly & Christenson, 2012).

Most educational researchers view school engagement as multidimensional (Fredricks et al., 2004) and the concept has been defined as "[a] student's active participation in academic and co-curricular or school-related activities, and commitment to educational goals and learning…. It is a multidimensional construct that consists of behavioral (including academic), cognitive, and affective subtypes" (Christenson et. al., 2012, pp. 816-817). This definition considers school engagement to be simultaneously manifested in active student participation, not only in behavioral terms, but also in emotional and cognitive terms (Li & Lerner, 2011). In other words, school engagement means whole-hearted active participation (see Dewey, 1913), as compared to mindlessly participating in class (being mentally absent; see Mosher & McGowan, 1985), or just having good intentions without actively participating in school-related activities.

The use of the term school engagement is most common in research on motivation (Li, 2011). The motivation literature is primarily focused on how students feel and think about things that can trigger action (Li, Lerner, & Lerner, 2010). Although some researchers use school engagement and

---

[1] In the research literature, authors use the terms school engagement and student engagement interchangeably (Libbey, 2004); thus, no distinction is made between these two terms in this study.

motivation as synonyms, they are two distinct constructs. Motivation is considered an internal cognitive and emotional state that can trigger action but does not include behavior, which is a key feature of school engagement (Li, 2011). Other studies, including the current study, consider school engagement to be a meta-construct that subsumes motivation in education (Fredricks et al., 2004; Martin, 2007).

Although school engagement is most commonly defined and studied in ways consistent with the discussion above, it should be pointed out that some scholars view school engagement as having social-behavioral components (Fredricks, Filsecker, & Lawson, 2016; Wang, Fredricks, Ye, Hofkens, & Linn, 2016) and others include aspects of self-regulation in their definition (Jarvela, Jarvenoja, Malmberg, Isohatala, & Sobocinski, 2016). As such, there is still confusion about what school engagement constitutes and the extent to which school engagement and ISR overlap within the field of educational motivation psychology (Boekaerts, 2016; Eccles, 2016). The current study contributes to an understanding of the construct by examining the nature of school engagement and its relation to ISR.

1.3.2   Measuring school engagement

At the outset of this study, few measures had included cognitive, emotional, and behavioral dimensions as parts of school engagement. Furthermore, few investigations had been devoted to evaluating the psychometric properties of school engagement measures and whether there was measurement invariance for different groups and ages (Li, 2011).

The current study conceptualized school engagement as the extent to which students are involved, connected, and committed to the academic and social activities provided in school (Li & Lerner, 2012). This definition of school engagement encompasses the three components of school engagement suggested by Fredricks et al. (2004) and Christenson et al. (2012) as described above. The operational definition used in the current study is called the Behavioral-Emotional-Cognitive School Engagement Scale (BEC-SES) and was developed by Li and Lerner (2011, 2013). The BEC-SES has been tested with students in Grades 9 through 11 within the 4-H Study in the U.S. and has shown evidence of both cross-group, as well as longitudinal, measurement equivalence (see Appendix C; Li & Lerner, 2012).

However, the high correlations frequently observed among behavioral, emotional, and cognitive engagement (see e.g., Li & Lerner, 2013) have raised questions about the multidimensionality of school engagement. Acknowledging the extent to which a measure is multidimensional is

important, as secondary dimensions can be lost when inappropriate models, such as one-factor models, or models that do not acknowledge the common variance of the factors, are fitted to multidimensional data (Ackerman, 1992; Reise, 2012).

Prior research on school engagement predicting academic achievement, using the correlated attributes model, suggested that only behavioral engagement strongly predicted academic achievement over emotional engagement and cognitive engagement (Chase, Hilliard, Geldhof, Warren, & Lerner, 2014; Li et al., 2010). Preliminary analysis in the current study revealed a discrepancy in the predictive value of school engagement for academic achievement depending on whether multidimensionality (i.e., the correlated attributes model) or the common variance (i.e., the unidimensional model) of school engagement was acknowledged in the statistical models that were fitted to the data.

The discrepancy observed in our preliminary analysis raised concerns about the validity of our measure of school engagement. As the validity of measures is a fundamental prerequisite of scientific research, we used a bifactor model to examine the extent to which our measure of school engagement (Li, 2011; Li & Lerner, 2013) was multidimensional or unidimensional (see Betts, 2012; Reise, 2012). Clarifying this conceptual and methodological issue was the principal aim of Paper II.

1.4 Paper III: School engagement and ISR: The potential to promote a positive integrated trajectory

The third and the last paper in this doctoral project addressed the overall goal of the study by testing the reciprocal relation between school engagement and ISR. Papers I and II supported this work by assessing the measures and analytical approach used in the Paper III. At this point in the dissertation work, my previous work on school engagement in Paper II had introduced me to theories originating in the field of educational motivational psychology that were relevant to my work and further supported the focus on the relation between ISR and school engagement. In the following section I present some of the literature that help formulate my argumentation for Paper III.

1.4.1 The model of motivational dynamics

As reviewed previously, the Five Cs Model of PYD (see Figure 1) highlights the importance of school engagement, ISR, and ecological assets for the promotion of PYD. An additional influential perspective on school

engagement is the self-determination theory perspective, which emphasizes the role of agency in human functioning (e.g., Jang, Kim, & Reeve, 2016; Reeve, Jang, Carrell, Jeon, & Barch, 2004; Skinner & Pitzer, 2012). The emphasis on a person's need for agency for the development of school engagement resonates strongly with the relational developmental perspective reviewed earlier, as both emphasize the importance of students' own activities and actions for increasing the probability of positive change. The model of motivational dynamics by Skinner and Pitzer (2012) stems from self-determination theory and is important for the current study, as the model defines school engagement and ISR as separate but related constructs, as does the previously reviewed PYD model. The reciprocal relation that is highlighted in the model of motivational dynamics underlines the potential to promote a positive integrated trajectory of school engagement and ISR. Testing this reciprocal relation was the main aim of Paper III and the overall aim of current study.

Unlike the Five Cs Model of PYD, the model of motivational dynamics is applicable specifically to the academic domain. The PYD covers a wide range of positive outcomes, while the model of motivational dynamics focuses on learning and achievement as outcomes. Similarly, the PYD model includes a wide range of ecological assets, while the model of motivational dynamics only specifies contextual support by teachers, parents, and peers. Both views are useful for the current study as they provide a framework to think about school engagement both within the classroom (model of motivational dynamics) and within the broader ecology of human development (the PYD model).

The model of motivational dynamics uses the concept "adaptive coping" to describe the ISR strategies that students use when faced with problems and difficulties related to schoolwork (Skinner & Pitzer, 2012). In the model, the reciprocal relation between school engagement and ISR is the closest antecedent of learning and achievement in a series of reciprocal feedback effects involving the student's context, the student's self-system processes, and the student's actions (see Figure 2).

According to Skinner and Pitzer (2012), school engagement is a sensitive indicator of the state of the motivational system as a whole. School engagement has been proposed to be very stable between and across school years during adolescence (see e.g., Gottfried, Fleming, & Gottfried, 2001; Skinner & Belmont, 1993). According to the model of motivational dynamics, the high stability of school engagement is due to the re-creation of stability by the feedback loops between engaged and disaffected actions,

on the one hand, and their dependents and antecedents, on the other.
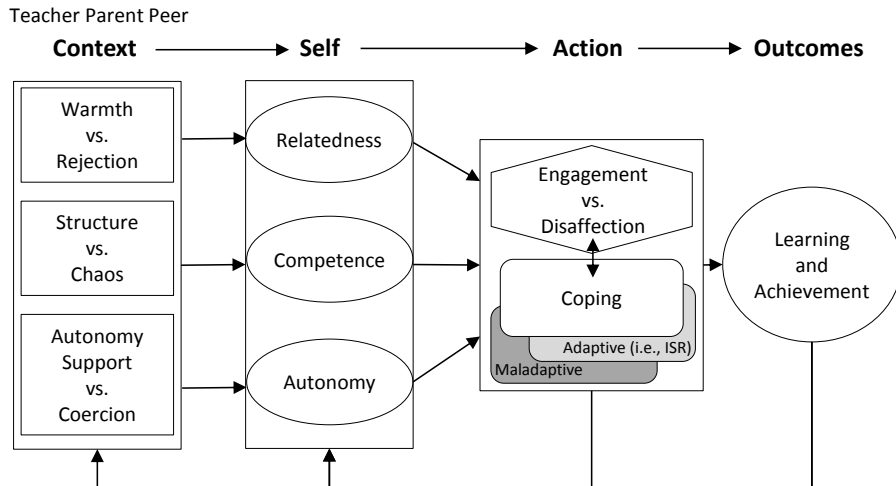
Figure 2. A dynamic model of motivational development organized around student engagement and disaffection. Figure adapted from Skinner and Pitzer (2012).

Students move through different school environments that each have different constellations of influences (e.g. teachers, peers, courses) that are, at some point, bound to affect their school engagement negatively (Eccles et al., 1993; Simmons, Burgeson, Carltonford, & Blyth, 1987). At these critical points in time, it must be important to have the means and flexibility (i.e., ISR) to protect school engagement from being lowered. Figure 2 highlights the expected reciprocal relations and the potential to promote a positive, integrated trajectory of school engagement and ISR.

At least two empirical studies (Berger & Karabenick, 2011; Ning & Downing, 2010) have assessed, but only provided partial support for, this relation (see Paper III). These two studies separately demonstrated predictive effects in opposite directions between ISR and school engagement, the studies failed to confirm the hypothesized reciprocal effects between the two constructs. However, both studies suggested that the lack of reciprocal relations might be due to specific circumstances related to the developmental periods being studied. Furthermore, research has underscored the importance of continuing efforts to investigate the hypothesized bidirectional relationship between school engagement and measures of ISR (Farley & Kim-Spoon, 2014; Karabenick & Zusho, 2015).

1.4.2   School engagement and ISR in context

The theoretical approaches (i.e., relational developmental systems, the PYD model, and the model of motivational dynamics) discussed in the previous sections, all emphasize that the development of adolescents is shaped by their own embodied activities and actions, which operate coactively in a lived world of physical and sociocultural objects. Subsequently, development leads to positive and negative feedback loops that are partly created by the adolescents own organized actions (Overton, 2015).

The contextual determinants of student's actions can be seen in the model of motivational dynamics (the left-most boxes in Figure 2) as warmth, structure and autonomy support provided by parents, teachers and peers. According to the model, these effects on student's actions are mediated through the self-beliefs of students. The importance of contextual determinants such as those depicted in Figure 2 are well supported by empirical research. Fredericks et. al. (2004) summarized that student engagement is higher in schools and classes that provide students with opportunities for voluntary choice, participation in school policy, and cooperation. Furthermore, engagement can be encouraged by clear and consistent goal setting, by limiting class size, and holding students accountable for deviant behavior. Support by teachers is an important facilitator of school engagement (Fredricks et al., 2004; Wang & Eccles, 2013). Teachers can provide warmth, structure and autonomy support by presenting relevant topics is an interesting way, setting challenging goals without overburdening students, and by providing formative feedback. The teachers expectations, stereotypes and communication with students also support or undermine engagement at school (Wang & Degol, 2013).

An important finding in the research on the contextual influences on school engagement is that contextual characteristics can influence school engagement differently depending on which subdimension of school engagement is being studied (Wang & Eccles, 2013). Furthermore, different context characteristics (e.g., peer values) can have contrasting effects on school engagement depending on the domain being studied (e.g., Math or English; see Leaper, Farkas, & Brown, 2012). These findings show that when thinking of the coaction of school engagement and the context, it is important to be explicit in terms of what people, domains, and subdimensions of school engagement are being studied to avoid findings being overgeneralized.

In the case of the current study, we studied students at the end of compulsory school in Iceland. The end of compulsory school is an important

time in Icelandic students' lives, as it has a direct bearing on their future academic prospects. Compared to earlier periods, tenth grade is characterized by an increasing urgency to achieve good grades (Sigthorsson, 2008). Students with high grades in Grade 10 can expect to be admitted to the most competitive upper secondary schools, whereas students with low grades have fewer schools to choose from and run a higher risk of dropping out of school (Blondal, Jonasson & Tannhäuser, 2011). Therefore, Grade 10 is generally viewed as a period of greater expectations and increased urgency with regard to educational goals. This context needs to be taken into consideration when interpreting the findings of the current study.

1.5 Principal aims of the study

In the previous sections, I have used three theoretical perspectives to frame a discussion about the importance of adolescent strengths (i.e., school engagement and ISR) for PYD, and for learning and achievement specifically. Furthermore, I used the same theories to argue for the reciprocal relation between school engagement and ISR in adolescence. Finally, I have identified several issues pertaining to the measurement of school engagement and ISR that need attention before valid conclusions can be made about the relations between the two constructs.

The goals of the current study therefore, were to address conceptual and methodological issues in the measurement of school engagement and ISR for use with adolescents in Iceland. The study also focused on the hypothesized reciprocal relation of school engagement and ISR, and the role of both constructs during the last two years of compulsory school in Iceland. The general and specific aims addressed by each of the three papers comprising my doctoral project, were as follows:

Paper I: The aim of the study was to address conceptual and methodological issues in the measurement of ISR. The specific aims were to:

1. Compare the reliability and validity of a forced-choice and a Likert-scale measure of ISR (i.e., the SOC measure).

2. Review the face validity of a Likert-scale measure of SOC.

3. Pilot a version of a Likert-scale measure of SOC with modified anchors.

4. Pretest a modified Likert-scale measure of SOC.

5. Confirm construct validity and longitudinal configural invariance of the SOC measure using confirmatory factor analysis (CFA).

Paper II: The aim of the study was to examine the validity of a multidimensional measure of school engagement in adolescence for use with Icelandic adolescents. The specific aims were to:

1. Examine the extent to which a measure of school engagement (i.e., the BEC-SES measure) was unidimensional vs. multidimensional.

2. Examine the criterion validity of BEC-SES by comparing how different representations (i.e., measurement models) of BEC-SES predicted academic achievement.

3. Test the best fitting BEC-SES model for configural, weak, and strong longitudinal factorial invariance.

Paper III: The aim of the study was to test the hypothesized positive reciprocal relation between school engagement and ISR during adolescence. The specific aims were to:

1. Test the configural, weak, and strong longitudinal factorial invariance of school engagement and ISR across four waves of measurement during Grade 9 and Grade 10.

2. Test whether a positive reciprocal relation existed between school engagement and ISR across the four times of measurement.

1.5.1   Significance of the study

The validation of measures of school engagement and ISR among adolescents in Iceland are of significance for Icelandic researchers and/or teachers who want valid measures to guide further research and development in the field of education. Furthermore, supporting the hypothesized reciprocal relation between school engagement and ISR is important as it highlights how various ISR-related strategies may promote school engagement through continued cycles of school engagement and ISR. For instance, through learning strategies (Zimmerman, 2002), approaches to maintain motivation in education (Wolters, 2003), adaptive help-seeking (Newman, 2002), or other knowledge-based pragmatics, such as practising mnemonic techniques to become a memory expert (Baltes & Baltes, 1990). Supporting the reciprocal relation between ISR and school engagement could prompt further examination of how these ISR-related strategies, and others, may co-develop with school engagement and, together, may support academic functioning.

2   Method

The Participants, Procedure, and Measures sections of the Methods chapter are organized by Paper I-III respectively. The Study design section is dedicated to the overall study design. Finally, to avoid repetition, the data analysis section addresses both the overall analytic techniques and analyses particular to each paper.

2.1   Participants

Five groups of students participated in the current study (see Table 3). Four of these five groups were cross-sectional samples that were only used for Paper I; group one: used for the first pilot (N = 139), group two: a focus-group (N = 15), group three: used for the second pilot (N = 42), and group four: used for pretest (N = 77). The fifth group, the main sample, was a four wave longitudinal sample (N = 561). Each group is described in detail in the order the data was collected in the following subsections.

2.1.1   Paper I: Group 1/First pilot

The participants in the first pilot, which compared Likert vs. forced-choice answer options on an ISR measure, consisted of a convenience sample of 139 undergraduate students in two research methodology courses at the University of Iceland that responded to an online questionnaire (response ratio 32%). The average age of the students was 29.6 years, with a standard deviation of 8.2 years, and 92% of the respondents were female.

2.1.2   Paper I: Group 2/Focus-group

The face-validity of the Likert-scale version of ISR was reviewed with a 15 student focus-group (33% girls; age 14) in a school in Reykjavik. The participants in these the focus groups were selected by convenience by the school principal.

2.1.3   Paper I: Group 3/Second pilot

The participants in the second pilot, examining new ISR answer options and the inclusion of ISR distractor items, were 42 Grade 9 students (50% girls; age 14) in a school in Reykjavik. The participants in the second pilot were selected by convenience by the school principal.

2.1.4   Paper I: Group 4/Pretest

The participants in the pretest, confirming the validity of a Likert version of SOC, were 77 Grade 9 students (47% girls; age 14) in a school in Reykjavik. The participants in the pretest were selected by convenience by the school principal.

2.1.5   Papers I through III: Group 5/Longitudinal sample

The fifth group of participants, the main sample, was a longitudinal sample (N = 561) of students in ninth grade at the beginning of the study in the fall of 2012 (most born in 1998). The participants in the main sample answered a paper survey four times at the beginning and end of Grade 9 and Grade 10. During the four waves of measurement, 539 (96%), 516 (92%), 519 (93%), and 510 (91%) participants returned a questionnaire, respectively. After the last wave, 81% of the participants had returned a questionnaire at all four waves of measurement.

When determining the sample size, the probability of making a false negative decision when evaluating the statistical significance of the coefficient of determination was set at 80%. Based on a review of previous related correlational research, the current study needed to have sufficient power to evaluate coefficients of determination as low as the .03 level with one regressor. The power estimate required that a minimum of 259 students were needed to participate to fulfill the given power requirements (Lenth, 2001).

To make the best use of the limited research funds available we limited the study to medium to large sized schools (>20 students in the ninth grade) in the Reykjavik area. Twenty of the 54 possible schools were selected randomly. Fifteen of the 20 schools agreed to participate. To ensure a sufficient group size at the school level, two ninth grade classrooms were randomly selected within each of the 15 schools. This method of sampling resulted in a nested data structure and introduced a bias towards students coming from medium sized schools. This bias however, was not considered problematic for the current study as the research questions were only at the individual level and not at the school level. However, to minimize the risk of making a Type 1 error when evaluating marginally significant effects in a sample with a nested data structure the COMPLEX feature of Mplus was used to produce correct standard errors using a sandwich estimator based on the school and class level clustering (see Muthén & Muthén, 1998-2012).

Table 3. The participants described in each research paper

	Paper I	Paper II	Paper III
Cross-sectional samples			
Pilot I (Undergraduates; N = 139)	X		
Focus-group (Grade 9; N = 15)	X		
Pilot II (Grade 9; N = 42)	X		
Pretest (Grade 9; N = 77)	X		
Longitudinal sample			
Wave 1 (Grade 9, fall; N = 539)	X	X	X
Wave 2 (Grade 9, spring; N = 516)	X	X	X
Wave 3 (Grade 10, fall; N = 519)		X	X
Wave 4 (Grade 10, spring; N = 510)		X	X

2.2 Procedure

2.2.1 Paper I: First pilot

The first pilot was conducted with the permission and help of the supervisor of two undergraduate methodology courses at the University of Iceland, School of Education. The supervisor notified the students of these courses about the upcoming survey and encouraged the students to participate. The survey was set up on survey system on the inner web of the University. The undergraduates were encouraged to comment on the measures in the survey by writing in a text box at the end of survey with the question "If you want to say something about the measures being tested, please write it in the text box below". Three email reminders were sent during a two week data collection period. As no personal information was collected, a formal consent was not considered necessary.

2.2.2 Paper I: Focus-group

The focus-group interview was conducted with the permission and help of a principal in Reykjavik. The researchers provided the principal with the necessary information about the study and the principal followed the schools protocol in notifying parents. As no personal information was collected, a signed parental consent was not considered necessary. The interview was conducted during school hours during a traditional 40 minute session. Two researchers were present and 15 students. The students

answered a Likert version of an ISR measure on paper and were urged to ask questions and comment on the measure they were answering. The researchers wrote down the comments made by the students. No other record was done during the focus-group interview.

2.2.3 Paper I: Second pilot

The second pilot was conducted with the permission and help of a principal in Reykjavik. The researchers provided the principal with the necessary information about the study and the principal followed the schools protocol in notifying parents. As no personal information was collected, a signed parental consent was not considered necessary. The measure was presented on paper. The data collection was conducted simultaneously in two classrooms during a single traditional 40 minute session. One researcher was present in each classroom. Each researcher read aloud a standardized instruction text. The classroom was organized the same way as when tests are administered.

2.2.4 Paper I: Pretest

The pretest was conducted with the permission and help of a principal in Reykjavik. The researchers provided the principal with the necessary information about the study and the principal followed the schools protocol in notifying parents. As no personal information was collected, a signed parental consent was not considered necessary. The survey was presented on paper. The data collection was conducted simultaneously in three classrooms during a single traditional 40 minutes session. Each researcher read aloud a standardized instruction text. The classroom was organized the same way as when tests are administered.

2.2.5 Paper I-III: Main survey

The main survey was conducted with the permission and help of principals and teachers in 15 schools in Reykjavik and neighboring municipalities. Signed parental consent was collected as personal information needed to be stored during the data collection period. The study was registered with the Personal Protection Authority with the registration number S5799. The research proposal was reviewed and approved during a formal interim evaluation at the School of Education carried out by an evaluation committee comprised of the supervisors and two external examiners.

The survey was presented on paper. The data collection was conducted simultaneously in two classrooms during a single traditional 40 minute

session. The classroom was organized the same way as when tests are administered. Standardized guidelines were created on how to present the survey in the classroom and graduate and undergraduate students were trained in administering the survey. One researcher was present in each classroom. These students and the Ph.D. candidate visited the 15 participating schools four times across three, six month intervals. Each participant was assigned a unique identifier on a sticker that was attached to his/her questionnaire. Graduate and undergraduate students entered the survey data. Personal information about the participants was stored on a password protected file server accessible only by the lead researchers, Kristján Ketill Stefánsson and Dr. Steinunn Gestsdóttir.

2.3   Measures

In the first paper of the doctoral project, I used three versions of an intentional self-regulation measure called SOC (Freund & Baltes, 2002). In addition, I used a measure of self-regulation called "Self-efficacy for self-regulated learning" (Bandura, 2006) to validate the most appropriate version of the SOC measure to use in the final paper of the doctoral project. In the second paper, I used a measure of school engagement called BEC-SES (Li & Lerner, 2011). In addition, I used a measure of academic achievement obtained from the Icelandic Educational Testing Institute to validate the most appropriate statistical model of BEC-SES to use in the final paper of the doctoral project. The measures used in the third, and the final, paper of the doctoral project consisted of the previously validated measures of SOC and BEC-SES together with a number of covariate measures (i.e., gender, socioeconomic status, mother's education, father's education, mother's occupation, father's occupation, age, school name, home language, and grade; OECD, 2012b).

For each scale measure in the final questionnaire, the model-based reliability estimate coefficient ω (Mcdonald, 1999) was calculated to indicate the proportion of the scale variance that was due to all common factors (Zinbarg, Revelle, Yovel, & Li, 2005). Coefficient ω is analogous to coefficient α (Reise, 2012); therefore, reliability estimates above the .70 level were interpreted as indicators of adequate reliability (Kline, 2011). All the measures used in the doctoral project are described in detail, by paper, below.

2.3.1 Paper I: Intentional self-regulation (forced-choice version)

During the piloting phase of the ISR measure, ISR was operationalized by using a short 24-item version of the SOC questionnaire (Freund & Baltes, 2002). This forced-choice version of the SOC measure included 24 forced-choice items where the respondent selects if he or she is more similar to Person A or Person B. The description of Persons A or B describes a self-regulated behavior (target items) or a non-self-regulated behavior (distractor items), respectively. This version of the SOC measure includes four subscales; Elective Selection, Loss-based selection, Optimization, and Compensation, each consisting of 6 items. This 24 item version of the SOC measure has been used in research with adolescents in the U.S. and has shown low reliability of the subscale but validity when used as a nine-item single factor construct (Gestsdottir & Lerner, 2007). The same measure has shown a poor fit for a three-factor structure and low reliability among older students (age 18) in Iceland (Gestsdottir et al., 2011). This measure was only used during the first pilot.

2.3.2 Paper I: Intentional self-regulation (first Likert version)

In accordance with results from an exploratory factor analysis on the forced-choice and Likert versions of the SOC measure (Geldhof, Bowers, et al., 2015; Geldhof et al., 2012) I adapted an Icelandic SOC version with five Likert answer options labelled "*Mjög miklu leyti*" (e. *Very Much*); "*Miklu leyti*" (e. *Much*); "*Svolitlu leyti*" (e. *Somewhat*), "*Litlu leyti*" (e. *Little*); "*Alls engu leyti*" (e. *Not at All*). The measure consisted of the 24 items used to indicate self-regulated behavior (target items) in the forced-choice version described earlier. This measure was only used during the first pilot and the focus group.

2.3.3 Paper I: Intentional self-regulation (second Likert version)

After the first pilot and the following focus group I adapted a second Likert scale version of SOC with new answer options. The new answer options were "*Mjög líkt mér*" (e. *Just like me*); "*Frekar líkt mér*" (e. *Somewhat like me*); "*Hvorki líkt né ólíkt mér*" (e. *Neither like me or not like me*); "*Frekar ólíkt mér*" (e. *Not like me*) og "*Mjög ólíkt mér*" (e. *Not at all like me*). The measure consisted of 18 items used to indicate self-regulated behavior (target items) and six items used to indicate non-self-regulated behavior (distractor items; a total of 24 items) in the forced-choice version described earlier. This was done to examine the possible effects of reverse-coded distractor items on the scales. This measure was only used during the second pilot.

2.3.4 Paper I: Self-efficacy for self-regulated learning (SRL)

A measure of self-efficacy for self-regulated learning (SRL) was adapted for use with adolescents in Iceland according to guidelines provided by Bandura (2006). A double blind translation process was used to translate the SRL measure to Icelandic. The adapted scale consisted of seven items. The respondents were asked to rate on a five point scale how confident they are that they can do each of the seven items. The five point scale ranged from "*cannot do*" (1) to "*highly certain can do*" (5). A sample item is "Get myself to study when there are other interesting things to do". The measurement of SRL has proven internally consistent with α coefficients ranging from .78 of .84 in previous studies (Usher & Pajares, 2008). The SRL measure was used during the first pilot, pretest, and during the analysis of data from Wave 1. Coefficient ω for the SRL measure at Wave 1 was .89. Appendix A contains a list of the scale items.

2.3.5 Paper I-III: Intentional self-regulation (ISR)

After two pilots, focus-group and pretest (see Paper I) intentional self-regulation (ISR) was operationalized by using a nine-item version of the SOC questionnaire (Freund & Baltes, 2002). The SOC questionnaire was originally translated into Icelandic by researchers with an extensive knowledge about the conceptual definition of SOC (Gestsdottir et al., 2011). Researchers in Iceland and the U.S. have used the nine-item version of SOC to measure a general ISR skill among adolescents, as discussed previously. The single adaptive ISR skill is manifested in nine items that each describes one of the sub-processes of SOC namely selection, optimization, and compensation. The respondents were asked how they decide what is important for them in life and how they go about achieving their goals in life. Following the question, the respondent had to indicate how well a list of statements adhered to the students behavior and/or cognitions on a five point scale ranging from "*Just like me*" (5) to "*Not at all like me*" (1; see Appendix B). A sample item is "I make every effort to achieve a given goal". Coefficient ω for the ISR measure used in Paper III was .71, .77, .77, and .82, respectively, by wave.

2.3.6 Paper II: Icelandic national examination

An assessment of academic achievement was retrieved from the Icelandic Educational Testing Institute. The institute conducts standardized achievement tests in fourth, seventh, and tenth grade every fall. Achievement data from Grade 10 was used in the current research. In

Grade 10 students are tested in English, Icelandic, and mathematics. The results are given on a standardized scale that ranges from 0 – 60 with an average of 30, and a standard deviation of 10 (Icelandic Educational Testing Institute, 2014).

2.3.7 Paper II-III: School engagement

In the current study, the Behavioral-Emotional-Cognitive School Engagement Scale (BEC-SES), developed by Li and Lerner (2011, 2013), was used to measure school engagement. The BEC-SES encompasses the three components of school engagement suggested by Fredricks et al. (2004); behavioral, emotional, and cognitive. Each component has five items (see Appendix C). The measure has been developed and tested with students in Grades 9 through 11 in the 4-H Study in the U.S. and has shown evidence of cross-group and longitudinal measurement equivalence (Li & Lerner, 2012). A double blind translation process was used to translate the measure of school engagement to Icelandic. After the translation, the measure was piloted, and finally pretested. During the pilot, the item "How often do you skip classes without permission?" showed high positive skew and considerable kurtosis (SI = 2.5; KI = 4.4; Kline, 2011). In an attempt to correct the high positive skew and kurtosis the anchor "*always*" was reworded to "*almost always*" during the pretest. During the pretest the skewness and kurtosis were lower (SI = 2.1; KI = 3.4) compared to the pilot so the change in translation was retained. During all four waves, the respondents were asked to rate on a four-point scale how often they do some of the five behavioral statements and how much they agree with the ten cognitive and emotional statements (see Appendix C). The four-point scale in the behavioral component ranged from "*Never*" (1) to "*Almost always*" (4) and the four point scale in the emotional and cognitive components ranged from "*Strongly disagree*" (1) to "*Strongly agree*" (4). Coefficient ω for the school engagement measure used in Paper III was .74, .80, .75, and .75, respectively, by wave.

2.3.8 Paper II-III: Background variables

Participants were asked for information about several background variables: gender, socioeconomic status, mother's education, father's education, mother's occupation, father's occupation, age, school name, home language, and grade. These background questions have been used in large scale international surveys and have shown good reliability among Icelandic youth (OECD, 2012b).

2.4 Study design

The study was a part of a larger research project on the positive development of youth in Iceland, directed by Dr. Steinunn Gestsdóttir. Prior to the main data collection, two pilot studies and one pretest were conducted where several hypotheses regarding the reliability and validity of the measurement ISR and school engagement were tested. The main data collection consisted of four waves of repeated group measurement with six month intervals (see Figure 3).
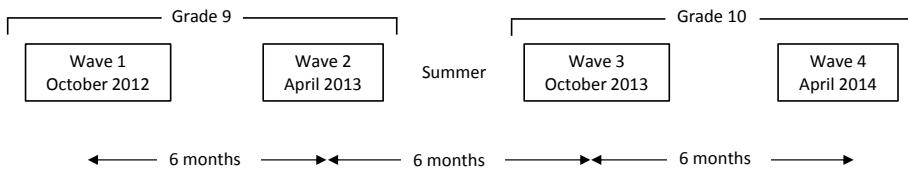
Figure 3. The main data collection consisted of four waves of repeated group measurement with six month intervals.

The first wave took place in October 2012, the second in April 2013, third wave took place in October 2013 and the fourth and final wave took place in April 2014, at the end of compulsory school. After the third wave of data collection, data from the Icelandic Educational Testing Institute was merged with the overall data set, providing an indicator of academic achievement in Grade 10. Data collection took place at the beginning and end of Grade 9 and Grade 10 as the study sought to capture changes in school engagement and ISR both within- and between-years towards the end of compulsory school.

2.5 Data analysis

In all the research papers, a *p*-value of ≤ 0.05 was considered significant. However, the significance level was occasionally lowered (see Paper III) to reduce the risk of making a Type 1 error due to high power in the longitudinal SEM analysis (N = 561; 4 waves). Model fit in factor analyses and structural equation models was, in all three papers, estimated by evaluating several fit indices: the chi-square statistic, the root mean square error of approximation (RMSEA), the comparative fit index (CFI), and the standardized root mean square residual (SRMR). Smaller chi-square, SRMR (SRMR ≤ .08), and RMSEA values (RMSEA ≤ .06), and higher CFI values (≥ .95) indicated a good model fit (West, Taylor, & Wu, 2013). Measurement invariance constraints in papers II and III were evaluated using a guideline

made by Cheung and Rensvold (2002), where a change of more than .01 in the comparative fit index (CFI) indicated that the assumption of invariance does not hold. Methods particular to each paper are addressed in the following paragraphs.

2.5.1 Paper I

The overall aim of the study presented in Paper I was to address measurement issues of ISR. The research hypotheses were tested by calculating descriptive statistics, reliability, and correlation coefficients using the SPSS 20 software package (IBM Corp., 2011) and conducting a series of factor analyses using version 7.1 of the Mplus software package (Muthén & Muthén, 1998-2012). The reliability of the ISR measure was evaluated by using the Cronbach's alpha reliability coefficient. A Cronbach's alpha coefficient of .7 - .8 was considered adequate (Kline, 1999). Convergent validity was evaluated by using the Pearson product-moment correlation coefficient and considered a coefficient up to .85 indicative of convergent validity. A correlation coefficient higher the .85 was considered indicative of two measures measuring the same construct (Kline, 2011). Descriptive statistics revealed that data collected by the final version of the ISR measure was normally distributed (see Paper I); therefore the latent factor models were estimated using the maximum likelihood estimation method. The estimates of latent factors were scaled using the marker variable method of scaling (see Little, 2013).

2.5.2 Paper II

The overall aim of the study was to examine the validity of a multidimensional measure of school engagement in adolescence for use with Icelandic adolescents. The research hypotheses were tested by estimating series of factor analyses and structural equation models using version 7.3 of the Mplus software package (Muthén & Muthén, 1998-2012). The estimates of latent factors were scaled using the fixed factor method (see Little, 2013), setting the variance of each latent factor to unity. A bifactor model was defined and indicated each specific factor by the items suggested by the previously established three-factor model (see Li & Lerner, 2012). In addition, a global school engagement factor was defined by all the items across the three specific factors. No cross-loadings or item-correlations were allowed. Finally, for identification purposes of the bifactor model, the correlations between all latent factors (general and specific) were set to zero within and across measurements (Reise, 2012).

Configural, weak, and strong longitudinal factorial invariance for the bifactor model was established using a method for models with ordered-categorical data described by Millsap and Yun-Tein (2004). Correlational analysis revealed significant correlations between several variables, such as self-reported grades, mother's education, father's education, and mother's occupation and missing cases at later waves. Accordingly, missing data were considered to be missing at random (MAR; see Little, 2013). These background variables were used to inform the creation of 20 imputed datasets without missing values using the multiple imputation feature of Mplus.

2.5.3   Paper III

The research hypotheses for Paper III were tested by estimating series of factor analyses and structural equation models using version 7.3 of the Mplus software package (Muthén & Muthén, 1998-2012). To reduce model complexity items of each of the constructs subdimensions were aggregated, using a method called parceling (see e.g., Bowers et al., 2016; Little, 2013). All measures for ISR and school engagement were treated as continuous variables. Furthermore, the latent factors were estimated and scaled using the fixed factor method of scaling. All models were fit to the data using the maximum likelihood estimator (see Little, 2013).

Next, a series of structural equation models with varying constraints were fitted to the data. In order to examine the theorized reciprocal effects, a longitudinal cross-lagged panel model was built by starting with freely estimating a minimal set of paths (the bivariate simplex process). Next, paths were added by evaluating information obtained from both modification indices and theory. For each step a likelihood ratio test (LRT) was conducted to compare the goodness of fit of the competing models. To avoid making a Type 1 error due to the high power of the longitudinal SEM model, a p-value less than .001 was chosen to determine a significant difference between competing models (see Little, 2013). Correlational analysis indicated that missing data was missing at random (MAR) and the full-information maximum likelihood estimation method (FIML) was used to handle missing data.

3 Results

The results presented in this section follow the order of the three papers that comprise this doctoral thesis. First, the development of the measure of ISR from Paper I is described in in detail in English, as the findings from Paper I are currently only available in Icelandic. Second, Paper II describes the development of the school engagement measure. Finally, Paper III presents the research findings on the reciprocal relations between school engagement and ISR.

3.1 Paper I: Addressing issues regarding the measurement of ISR

The first paper was based on data collected in two pilot studies and one pretest where several hypotheses regarding the reliability and validity of the measurement of ISR (i.e., the SOC measure) in adolescence were tested. As previous research has found the SOC measure to have problematic psychometric qualities when used with U.S. and Icelandic youth (Geldhof, Bowers, et al., 2015; Geldhof, Gestsdottir, et al., 2015; Geldhof et al., 2012), various steps were taken to ensure the validity and reliability of the measure.

The SOC measure I started with in the current study included 24 forced-choice items where the respondent selects if he or she is more similar to Person A or Person B. The description of Persons A or B describes a self-regulated behavior or a non-self-regulated behavior, respectively. This version of the measure included four subscales; Elective Selection, Loss-based selection, Optimization, and Compensation, each consisting of 6 items (Baltes et al., 1999). This 24 item version of the SOC measure has been used in research with adolescents in the U.S. and has shown low reliability of the subscale but validity when used as a nine-item single factor construct (Gestsdottir & Lerner, 2007). The low measurement reliability laid the grounds for the first empirical question of Paper I: Will Likert-scale answer options make the SOC measurement more reliable and valid as suggested by Geldhof et al. (2012)?

3.1.1 First pilot: Comparing Likert vs. forced-choice answer options

To compare the Likert-scale options and the forced-choice scale, a convenience sample of 139 undergraduate students answered an online questionnaire. At the end of the survey, the students were urged to comment on the items and scales being tested. Twenty-one students commented on the questionnaire and 15 of them indicated that the forced-choice version was confusing and hard to answer and that the Likert-scale version was easier to understand and answer. The remaining six students had comments on other items in the questionnaire or gave answers that were not relevant to scale development. The students' opinions expressed in the open-ended questions was supported by a reliability analysis that showed that the reliability of the SOC subscales moved from being poor in the forced-choice version to being acceptable or close to acceptable using the Likert version (see Table 4).

Table 4. Reliability coefficients of SOC (forced-choice and Likert-scale versions) from a pilot conducted with undergraduate students (*N* = 139)

	Cronbach's alpha SOC (forced-choice version)	Cronbach's alpha SOC (Likert-scale version)
Selection	.49	.71
Optimization	.56	.60
Compensation	.39	.66
Loss based selection	.55	.74
SOC (9 item version)	.67	.77

Both versions of the SOC measure showed significant correlation to the theoretically related measure of self-efficacy for self-regulated learning (SRL). The strength of the relationship was higher using the Likert version of SOC rather than the forced-choice version (see Table 5).

Table 5. Pearson's correlation of SOC (forced-choice and Likert-scale versions) with self-efficacy for self-regulated learning (SRL) from a pilot conducted with undergraduate students (*N* = 139)

	Correlation of SOC (forced-choice version) with SRL	Correlation of SOC (Likert-scale version) with SRL
Elective selection	.20*	.32**
Optimization	.25**	.42**
Compensation	.23**	.32**
Loss based selection	.28**	.29**
SOC (9 item version)	.43**	.49**

The skewness of the scales was closer to normal distribution in the Likert-scale version. The kurtosis was not good in the Likert-scales of Optimization and SOC (Table 6).

Table 6. Skewness and distribution of SOC (forced-choice and Likert versions) from a pilot conducted with undergraduate students (*N* = 139)

	Skewness		Kurtosis	
	SOC (forced-choice version)	SOC (Likert-scale version)	SOC (forced-choice version)	SOC (Likert-scale version)
Elective selection	0.12	0.12	-0.57	-0.57
Optimization	-0.77	0.2	0.1	**1.1**
Compensation	-0.45	-0.2	-0.35	0.12
Loss based selection	-0.84	0.2	-0.16	0.74
SOC (9 item version)	-0.55	0.2	-0.28	**1.31**

Note. Poor values are bolded.

3.1.2 Focus-group: Inspecting the face validity of a Likert-scale version of SOC

Considering the difficulties adults had with answering the forced-choice versions of SOC, and the results indicating that the Likert-scale version of SOC was equally or more valid than the forced-choice version, the Likert-scale version was chosen to be used. Next, the face-validity of the Likert-scale version was reviewed with a 15 student focus-group (Age 14) in a school in Reykjavik. In the focus-group discussion, a few students pointed out that the answer options "*Not at all*," "Little," "*Somewhat*," "*Much*," and "*Very much*" was difficult to understand in Icelandic. The students recommended a more common set of Likert-scale answer options: "*Just like me*," "*Somewhat like me*," "*Neither like me or not like me*," "*Not like me*," and "*Not at all like me*." The answer options recommended by the focus-group turned out to be more appropriate when compared to the instructions to the original forced-choice version of the SOC questionnaire (see Baltes et al., 1999). In addition, at this point in the development a few of the ISR target items (25%) were replaced with the distractor items to examine possible effects of reverse-coded distractor items on the scales.

3.1.3 Second pilot: Examining new answer options and the inclusion of distractor items

Another pilot study was conducted with 42 Grade 9 students, in another school in Reykjavik, to examine if the change of answer options from "*Not at all*," "*Little*," "*Somewhat*." etc. to "*Just like me*," "*Somewhat like me*," etc. would change distribution of the answers and the possible changes of including a few ISR distractor items in the Likert-scale version. The results indicated that the change in answer options had a beneficial effect on the distribution of the answers to the ISR target items compared to answers previously given by undergraduate students (see Table 6). These results were later supported in another pretest (see Table 9). However, the reliability coefficients (see Table 7) were unacceptable when including the ISR distractor items. Further analysis revealed that many of the ISR reverse coded distractor items showed negative inter-item correlations to other items in the scales. Therefore, after the second pilot study, the inclusion of ISR distractor items was not considered in the final Likert-scale version.

Table 7. Reliability coefficients of SOC (Likert version) including reverse coded ISR distractor items from a second pilot with 14 year old students (*N* = 42)

	Cronbach's alpha SOC (with ISR distractor items)
Elective selection	.49
Optimization	.35
Compensation	.28
Loss based selection	.48
SOC (9 item version)	.57

3.1.4 Pretest: Confirming the validity of a Likert version of SOC

After the focus-group and two rounds of pilot testing, a final version of the questionnaire was pretested with 14 year old students in a school in Reykjavik (*N* = 77). The final version included the SOC questionnaire with 24 ISR target items and the following Likert-scale answer options: "*Just like me*," "*Somewhat like me*," "*Neither like me or not like me*," "*Not like me*," and "*Not at all like me*." The results showed an acceptable reliability for the scales of Optimization and Loss-based selection. The Selection scale also had acceptable reliability by deleting the item: "I concentrate all my energy on few things." (see Table 8). The compensation scale still had questionable reliability in contrast with the undergraduate students tested previously, wherein the Compensation scale had acceptable reliability (see Table 4).

Table 8. Reliability of SOC (Likert version) with ISR target items only from a pretest with 14 year old students (*N* = 77)

	Cronbach's alpha SOC (Likert-scale version)	Cronbach's alpha if one item deleted	Items reducing reliability
Selection	.49	.64	**a) I concentrate all my energy on few things.**
Optimization	.71	.72	f) When I want to achieve something difficult, I wait for the right moment and the best opportunity.
Compensation	.44	.49	d) When something does not work as well as before, I get advice from experts or read books.
Loss based selection	.80	.81	s) When I can't do something as well as I used to, I think about what exactly is important to me.
SOC (9 item version)	.77	.78	r) I always pursue goals one after the other.

Note. Item improving alpha by more than .1 when deleted is bolded

Item analysis during the second pilot study indicated that the new Likert-scale answer options ("*Just like me*," "*Somewhat like me*"...etc.) yielded a distribution closer to normal than did the response options used in the first pilot study ("*Not at All*," "Little," "*Somewhat*"… etc.). This finding was supported in the pretest, as the previous kurtosis of the Optimization and SOC (nine-item version) Likert-scales in the first pilot study was reduced to a good value in the pretest (see Table 9).

Table 9. Skewness and kurtosis and correlation with self-efficacy for self-regulated learning from a pretest with 14 year old students (*N* = 77)

	Correlation of SOC (Likert version) with self-efficacy for self-regulated learning	Skewness	Kurtosis
Selection	.10	-0.13	0.15
Optimization	.45**	0.05	-0.78
Compensation	.18	0.03	0.29
Loss based selection	.21	-0.3	-0.7
SOC (9 item version)	.51**	0.15	-0.57

Note. **significant at the .01 level; * significant at the .05 level.

3.1.5 Conclusions from the development of the Icelandic version of the SOC measure

Based on the results of the two pilots, the focus-group, and the pretest, the final version included a 24 item version of the SOC questionnaire with ISR target items only, using the following answer options: "*Just like me*," "*Somewhat like me*," "*Neither like me or not like me*," "*Not like me*," and "*Not at all like me*." Consistent with previous studies with U.S. youth using SOC as a single nine-item construct, this measure seemed to give the most valid measure of ISR among Icelandic youth.

3.1.6 Confirmatory factor analysis of the Icelandic SOC measure

Data from Wave 1, in the main longitudinal study, provided sufficient power (*N* = 539) to perform a confirmatory factor analysis on the SOC measure (see Brown, 2015; Little, 2013). At first, a three-factor correlated attributes model was fit to the data as described by theory (Baltes et al., 1999). The three-factor model did not fit the data well (χ^2 = 574.72, *p* = 0.00; SRMR = .07; RMSEA = .08 (CI = .08-0.09); CFI =.82). On closer inspection, the correlation among the three factors was very high (*r* = .79 to .98). In light of the high correlations between the three factors a model with a single nine-item general SOC factor was tested instead of the three-factor model and plans for fitting a four factor model (including loss-based selection) were not considered further. The use of a nine-item general SOC factor in

adolescence is in accordance with recommendations from prior research in the U.S. (Gestsdottir et al., 2009; Zimmerman et al., 2007).

When the nine-item general SOC factor was fit to the data, the fit indices revealed a good fit (χ^2 = 70.03, *p* = .00; SRMR = .04; RMSEA = .06 (CI = .04-.07); CFI = .95). The reliability of the nine-item SOC factor was acceptable (α = .75). Finally, longitudinal configural invariance was confirmed by fitting the nine-item model to data from Wave 2 (χ^2 = 58.34, *p* = 0.00; SRMR = .03; RMSEA = .05 (CI = .031-.07); TLI = .96; CFI = .97). The model fit remained good at Wave 2. The reliability of the nine-item SOC factor at Wave 2 was also good (α = .84). A single factor structure using nine-items from the SOC questionnaire indicated a good fit, as has been found by previous research with adolescents in other cultures (see Table 10). These results indicated that a single factor, nine-item SOC measure with Likert answer options gave a reliable and valid measure of ISR among students at the beginning and end of ninth grade in Iceland.

Table 10. SOC items remaining after a confirmatory factor analyses among 14 and 18 year old participants in Iceland and 11-13 year old participants in the U.S. (adapted from Paper I)

		Iceland			U.S.		
		14	14*	18*	11#	12#	13+
Selection items							
m)	When I decide upon a goal, I stick to it	x	x	x	x	x	x
r)	I always pursue goals one after the other	x	x	x	x	x	x
a)	I concentrate all my energy on a few things						
b)	I consider exactly what is important for me						
k)	I always focus on the one most important goal at a given time						
l)	When I think about what I want in life, I commit myself to one or two important goals						
Optimization items							
c)	I keep trying until I succeed at a goal	x	x	x	x	x	x
f)	When I want to achieve something, I can wait for the right moment						
g)	I think about exactly how I can best realize my plans	x		x	x	x	x
h)	I make every effort to achieve a given goal	x	x	x	x	x	x
j)	When I have started something that is important to me, but has little chance at success, I make a particular effort	x	x	x	x	x	x
n)	When I want to get ahead, I take a successful person as a model						
Compensation items							
e)	For important things, I pay attention to whether I need to devote more time or effort	x		x	x	x	x
o)	When things don't work the way they used to, I look for other ways to achieve them	x	x	x	x	x	x
q)	When something doesn't work as well as usual, I look at how others do it	x	x	x	x	x	x
d)	When something does not work as well as before, I listen to advisory broadcasts and books as well						
i)	When things don't work the way they used to, I look for other ways to achieve them						
p)	When I can't do something as well as I used to, then I ask someone else to do it for me						

Note. The numbers refer to the age of the participants; *(Gestsdottir et al., 2011); #(Gestsdottir & Lerner, 2007); +(Zimmerman et al., 2007).

3.2 Paper II: Addressing issues regarding the measurement of school engagement

The aim of Paper II was to examine the validity of a multidimensional measure of school engagement in adolescence for use with Icelandic adolescents. School engagement involves cognitive, emotional, and behavioral components that overlap conceptually (Fredricks et al., 2004). This conceptual ambiguity has led to measures that have consisted either of one general factor or of separate correlated factors. However, neither approach can sufficiently account for both the uniqueness and the overlap of the subcomponents (Betts, 2012).

During the pretest, described in the previous section, the school engagement measure (Li & Lerner, 2013), indicated adequate reliability (α = .69 to .84) for all the school engagement subscales (i.e., behavioral engagement, emotional engagement, and cognitive engagement). However, during later analysis, a discrepancy in model results appeared. When the traditional correlated attributes model was used, only behavioral engagement predicted later academic achievement (β = .71). However, when a unidimensional or a hierarchical model was used the prediction went from very strong (β = .71) to strong (β = .51). Due to the discrepancy observed, it became necessary to determine, using a bifactor model, the degree to which the school engagement measure was unidimensional versus multidimensional. This work became the subject of the second paper of the Ph.D. project. The results of this inspection are summarized in the following section, for detailed results see Paper II.

3.2.1 Confirmatory factor analyses of the school engagement measure

The WLSMV estimation method was used to fit three measurement models to the data: a one-factor model, a three-factor model, and a bifactor model. Model identification was established by fixing the variance of each latent variable to unity. Model fits are in Table 11. The one-factor model exhibited inadequate fit (χ^2 (90) = 552.37; CFI = .90; RMSEA = .10) because of large chi-square and RMSEA values and a low CFI value. The three-factor model showed a good fit (χ^2 (87) = 227.68; CFI = .97; RMSEA = .05), with a significant reduction in the chi-square value compared to the nested one-factor model. In addition, the three-factor model showed an acceptable RMSEA value and a good CFI value. The bifactor model, however, provided the best fit of the three models, with the lowest chi-square value and good RMSEA and CFI values (χ^2 (75) = 149.89; CFI = .98; RMSEA = .04). A chi-

square difference test using the DIFFTEST option in Mplus confirmed that the three-factor model fit the data better than the nested one-factor model ($\Delta\chi^2(3) = 211.65$, $p < .001$), and that the bifactor model fit the data better than the three-factor model ($\Delta\chi^2(12) = 87.84$, $p < .001$).

Table 11. CFA fit statistics for the Behavioral-Emotional-Cognitive School Engagement Scale measurement models at Wave 1 (reproduced from Paper II)

	χ^2	s	df	$\Delta\chi^2$	Δdf	CFI	RMSEA
One-factor model	552.373	15.52	90			.90	.10
Three-factor model	227.675	8.17	87	324.70	3	.97	.05
Bifactor model	149.885	5.38	75	77.79	12	.98	.04

Note. These are average results over 20 data sets; Chi-square difference tests were conducted between nested models at Wave 1 with non-imputed data.

The standardized factor loadings of the different models are in Table 12. The one-factor model was well defined and highly reliable ($\omega = .93$), with factor loadings ranging from .42 (participation in classroom discussions) to .79 (caring about the school). The three-factor model was well defined and reliable (behavioral engagement, $\omega = .82$; emotional engagement, $\omega = .87$, and cognitive engagement, $\omega = .90$), with factor loadings ranging from 0.49 to 0.88. All the subfactors in the three-factor model correlated strongly with each other, with latent correlation coefficients ranging from $r = .65$ between emotional and cognitive engagement to $r = .72$ between cognitive and behavioral engagement. The remaining correlation between emotional and behavioral engagement was $r = .66$.

The general school engagement scale in the bifactor model was also well defined and highly reliable ($\omega = .93$), with factor loadings ranging from 0.40 (come to class unprepared) to 0.76 (learn as much as I can at school). Although also highly reliable (behavioral engagement, $\omega = .84$; emotional engagement, $\omega = .87$, and cognitive engagement, $\omega = .91$), the three specific factors were less well defined than the general factor. All the specific factor loadings were significant at the $p < .01$ level, although one behavior engagement factor item, which refers to participation in class discussions (Item 4), showed a particularly low loading (see Table 12 and Appendix C).

Table 12. **Standardized factor loadings of the three measurement models for the Behavioral-Emotional-Cognitive School Engagement Scale at Wave 1 (adapted from Paper II)**

Nr.	One-factor SES	Three-factor B	Three-factor E	Three-factor C	Bifactor G-SES	Bifactor S-B	Bifactor S-E	Bifactor S-C
1	0.48**	0.56**			0.40**	0.82**		
2	0.70**	0.80**			0.67**	0.46**		
3	0.67**	0.74**			0.64**	0.30**		
4	0.42**	0.49**			0.44**	0.11**		
5	0.70**	0.85**			0.74**	0.20**		
6	0.74**		0.81**		0.64**		0.45**	
7	0.79**		0.88**		0.64**		0.67**	
8	0.75**		0.83**		0.62**		0.56**	
9	0.55**		0.62**		0.47**		0.41**	
10	0.58**		0.66**		0.56**		0.29**	
11	0.75**			0.81**	0.76**			0.21**
12	0.75**			0.85**	0.74**			0.39**
13	0.73**			0.76**	0.60**			0.51**
14	0.72**			0.77**	0.64**			0.47**
15	0.74**			0.80**	0.60**			0.67**

$**p < 0.01$; *Note.* These are average results over 20 data sets; SES = School Engagement Scale; B = Behavioral; E = Emotional; C = Cognitive; G-SES = General School Engagement Scale; S-B = Specific behavioral; S-E = Specific emotional; S-C = Specific cognitive.

3.2.2 Criterion validity: Latent regression analyses of school engagement and academic achievement

Icelandic national examination scores (INE) were added at Wave 3, as a continuous outcome variable, to the three-factor and the bifactor measurement models from Wave 1 to assess the relative performance of the different measurement models in predicting academic achievement (see Figure 4). The one-factor model was not included due to the poor model fit established in the CFA (see Table 11).

The WLSMV estimation method was used to fit these structural equation models to the data, and model identification was enabled by setting the variance of each latent variable to unity. The fit indices of the models and latent regressions are shown in Table 13. The three-factor and the bifactor models showed a good fit; the bifactor model at Wave 1 fit the data significantly better than the three-factor model according to a chi-square difference test using the DIFFTEST option in Mplus ($\Delta\chi^2(13) = 61.25$, $p < .001$).
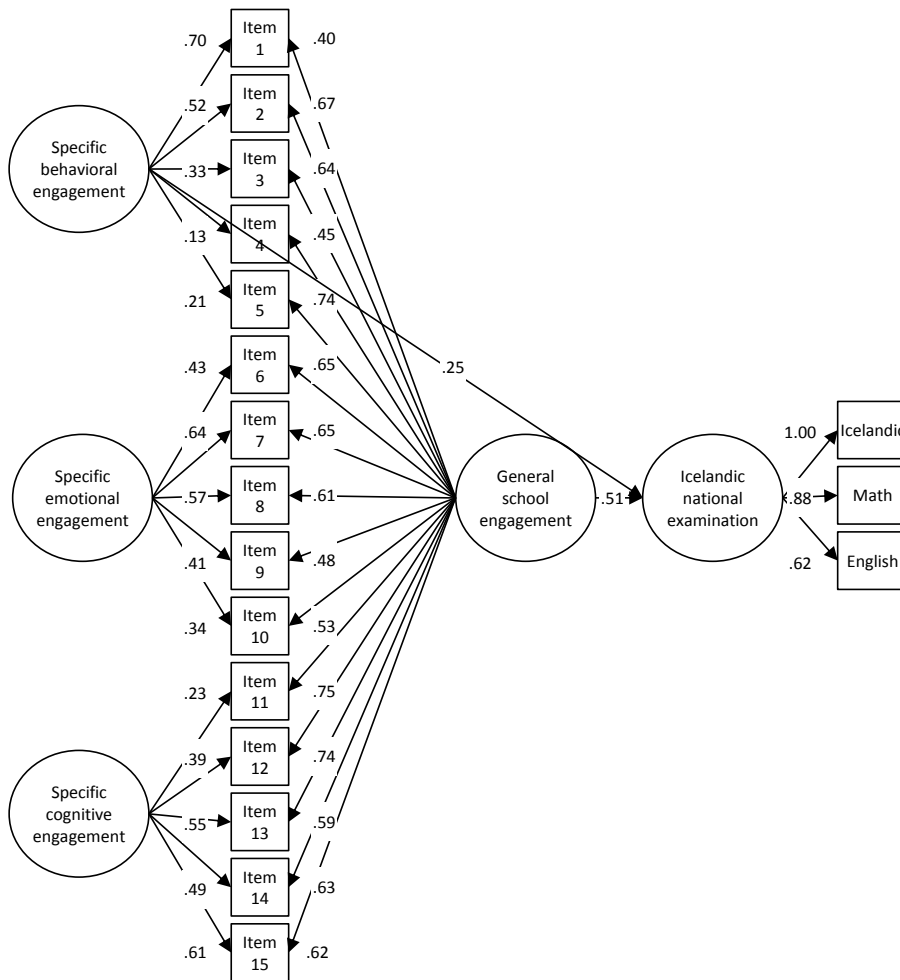
Figure 4. Empirical results of a structural equation model where a bifactor model of school engagement at the beginning of Grade 9 (Wave 1) predicts Icelandic national examination scores at the beginning of Grade 10 (Wave 3). Total number of participants = 561. The variances of the latent factors were set to unity to allow for identification. For clarity, only significant (*p* < .01) factor loadings and regression coefficients are shown in the diagram. Non-significant regression coefficients and fit indices can be found in Table 13. Adapted from Paper II.

Table 13. SEM fit statistics and standardized regression coefficients of the school engagement measurement models at Wave 1 predicting Icelandic national examinations scores (Icelandic, mathematics, and English combined) at Wave 3 (reproduced from Paper II)

	χ^2	s	df	CFI	RMSEA	R^2	β
Three-factor model	263.462	7.09	129	.97	.04	.36	
Behavioral engagement							0.73**
Emotional engagement							-0.09
Cognitive engagement							-0.12
Bifactor model	204.369	5.21	116	.98	.04	.36	
General school engagement							0.51**
Specific behavioral engagement							0.25**
Specific emotional engagement							-.12
Specific cognitive engagement							-.13

**$p < 0.01$; *$p < .05$; Note. These are average results over 20 data sets; Chi-square difference tests between nested models at Wave 1 were conducted with non-imputed data.

3.2.3 Factorial invariance of the bifactor model of school engagement

Finally, in order to ensure that the structure of school engagement did not substantially vary over time, the last analytic step was to test factorial invariance of the bifactor solution, the consistency of measurement of the bifactor model was examined by testing configural, weak, and strong factorial invariance across the four waves of available data. Scale identification was obtained by using guidelines described by Millsap and Yun-Tein (2004), the results can be seen in Table 14. The configural invariance model showed excellent fit with an average CFI of .978 and a standard deviation of only .001 across the 20 datasets. The weak invariance model was specified by fixing the individual factor loadings to be equal across the four waves. This specification caused a very small improvement in model fit, increasing the CFI by .001 while the standard deviation of the CFI remained small (.001). The strong invariance model was further specified by fixing individual thresholds to be equal across the four waves. The strong invariance model gave the same CFI and standard deviation as the weak invariance model. Differences in CFI between invariance models were well below the .01 criterion chosen for the comparison, which

supported configural, weak, and strong factorial invariance across the four waves.

Table 14. Model fit statistics for the tests of measurement invariance of general and specific aspects of behavioral, emotional and cognitive engagement across four waves (reproduced from Paper II)

	χ^2	s	df	$\Delta\chi^2$	Δdf	CFI	s	RMSEA	s
Configural model	1843.028	9.43	1536			.978	.001	.019	.000
Weak invariance	1909.856	11.25	1614	66.83	78	.979	.001	.018	.000
Strong invariance	1987.091	11.49	1692	77.24	78	.979	.001	.018	.000

Note. These are average results over 20 data sets.

In the three-factor model, only the behavioral engagement factor strongly predicted subsequent INE scores (β = 0.73, 95% CI [0.49, 0.98]). In contrast, the bifactor model at Wave 1 produced two separate direct effects. The general school engagement factor produced a strong direct effect (β = 0.51, 95% CI [0.37, 0.65]) and, in addition, the specific behavioral engagement factor produced a moderate direct effect (β = 0.25, 95% CI [0.06, 0.44]) on the INE scores. The specific emotional and specific cognitive factors had weak and non-significant effects. The bifactor and three-factor school engagement models at Wave 1 both explained 36% of the variance of the INE scores.

3.3 Paper III: The reciprocal relation between school engagement and ISR

Based on the relational developmental systems metatheory, the 5 Cs model of PYD (see Figure 1) and the model of motivational dynamics (see Figure 2) I have argued for the importance of, and bidirectional relation between school engagement and ISR. More specifically, I argued for the reciprocal effects between school engagement and ISR during the end of compulsory school in Iceland. This theoretical argument was supported with empirical data in Paper III by testing the hypothesis that a positive reciprocal relation existed between school engagement and ISR across four waves of data with students in Grades 9 through 10.

3.3.1   Longitudinal invariance of school engagement and ISR

In order to ensure that the same latent construct was being measured at each measurement occasion, the maximum likelihood estimation method was used to fit a series of bivariate longitudinal measurement models with varying constraints to the data. The residual variances of the corresponding indicators were allowed to correlate over time and estimates of latent factors were scaled using the fixed factor method of scaling. The configural model gave an excellent fit to the data, with CFI (.99) and RMSEA (.026). Next, each factor loading was constrained to be equal across time, these constraints led to a minor decrease in model fit (ΔCFI = .003), suggesting weak factorial invariance. In addition, each intercept was constrained to be equal across time. Again, these constraints led to a minor decrease in model fit (ΔCFI = .005), suggesting strong factorial variance for the bivariate longitudinal measurement models (see Table 15).

Table 15. Model fit statistics for the tests of measurement invariance of ISR and school engagement across four waves (reproduced from Paper III)

	χ^2	df	p	RMSEA	RMSEA 90% CI	CFI	ΔCFI
Null model	7196.109	312	<.001	---	---	---	---
Configural invariance	259.337	188	<.001	.026	.018;.033	.990	---
Weak invariance	288.789	200	<.001	.028	.021;.035	.987	.003
Strong invariance	334.242	212	<.001	.032	.025;.038	.982	.005

3.3.2   Descriptive estimates of the latent variables

Latent variable estimates derived from the well-fitting strong invariance model revealed small mean differences in school engagement (d_{max} = 0.14) and very small differences in ISR (d_{max} = 0.05) across the four waves of measurement (see Table 16). An omnibus test of the latent means indicated marginally significant differences ($\Delta\chi^2$(5) = 19.20, p = .002) of school engagement and ISR across the four waves measured. In general, school engagement tended to be slightly lower during the end of the school year compared to the beginning of the school year (see Table 2).

Table 16. Standardized latent means for school engagement (SES) and ISR measured at the beginning and the end of Grade 9 (Waves 1 and 2) and Grade 10 (Waves 3 and 4; reproduced from Paper III)

	Wave 1		Wave 2		Wave 3		Wave 4	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
ISR	0.00	1	0.05	1.18	0.02	1.19	0.03	1.31
SES	0.00	1	-0.14	1.18	-0.01	1.06	-0.11	1.13

Note: ISR = Intentional self-regulation; SES = School engagement scale.

Correlational estimates (see Table 17) revealed very strong correlations (*r* ranged from .78 to .82) between measures of school engagement at successive time points. Similarly, the correlations between ISR at successive time points were strong (.63 to .71). Furthermore, the analysis revealed strong correlations between school engagement and ISR within each measurement occasion (.56 to .67), as well as strong correlations between school engagement and ISR across consecutive measurement occasions (.50 to .66).

In sum, the latent means and correlations indicated that the measures of school engagement and ISR were assessing the two constructs in a reliable way across time. In addition, the strong within- and between-construct associations indicated the potential of detecting cross-lagged effects by fitting a structural model to the data.

Table 17. Latent bivariate correlations between school engagement and ISR derived from the baseline model (strong invariance) across waves (reproduced from Paper III)

			ISR				SES			
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ISR	Wave 1	(1)	-							
	Wave 2	(2)	0.70	-						
	Wave 3	(3)	0.61	0.71	-					
	Wave 4	(4)	0.53	0.66	0.63	-				
SES	Wave 1	(5)	0.64	0.59	0.51	0.42	-			
	Wave 2	(6)	0.55	0.63	0.59	0.43	0.82	-		
	Wave 3	(7)	0.54	0.58	0.67	0.50	0.75	0.78	-	
	Wave 4	(8)	0.50	0.54	0.66	0.56	0.69	0.75	0.79	-

Note: N = 561; All correlations significant at *p* < .001 level; ISR = Intentional self-regulation; SES = School engagement scale.

3.3.3 Longitudinal structural models

After establishing strong factorial invariance, the strong invariance measurement model was used as a baseline model to further test a series of longitudinal structural equation models in a nested hierarchical manner (see Little, 2013). First, a model of the bivariate simplex process was fit to the data. This first model gave a good fit to the data (see Table 18) but the fit was significantly worse than the baseline model ($\Delta\chi^2(18)$ = 126.48, *p* < .001). The addition of contextual effects, by allowing measures at Wave 1 to predict measures at Wave 3, and by allowing measures at Wave 2 to predict measures at Wave 4 (i.e., school start/end effects; see Little, 2013), improved the model fit but the model still fitted the data worse than the baseline model ($\Delta\chi^2(14)$ = 58.21, *p* < .001). The addition of partial cross-lagged effects (i.e., ISR predicting school engagement) further improved the model fit and the model was no longer significantly worse fitting than the baseline model ($\Delta\chi^2(11)$ = 24.73, *p* = .01). Next, the full cross-lagged model was fitted to the data, this model did not significantly reduce the model fit compared to the baseline model ($\Delta\chi^2(8)$ = 8.62, *p* = .38), and gave the best fit to the data of all the longitudinal structural models (χ^2 (220) = 342.87; RMSEA = .04; CFI = .98). In addition, the full cross-lagged model was compared to a constrained model where the cross-lagged paths were constrained to be equal across time (see e.g., Engels et al., 2016). The

constrained cross-lagged model did not fit the data significantly worse than the full cross-lagged model ($\Delta\chi^2(4) = 5.61$, $p = .23$) and was retained as the most appropriate model to answer the research question.

Table 18. Summary results from the bivariate model building process (reproduced from Paper III)

	χ^2	df	$\Delta\chi^2$	Δdf	p	RMSEA	CFI
Baseline (strong invariance)	334.242*	212	-	-	-	.032	.982
Bivariate simplex	460.724*	230	126.482	18	<.001	.042	.966
Contextual	392.456*	226	58.214	14	<.001	.036	.976
Partial cross-lagged	358.967*	223	24.725	11	.010	.033	.980
Full unconstrained cross-lagged	342.866*	220	8.624	8	.375	.032	.982
Full constrained cross-lagged	348.474*	224	5.608	4	.230	.032	.982
Final model with covariates	556.271*	315	-	-	-	.037	.966

Note: $N = 561$; * = significant at $p < .001$ level.

Finally, a set of common covariates (i.e., gender, socio-economic status, parents education and occupation, and self-reported grades) were included in the final model to avoid confounding the relations between the main study variables. The covariates were treated as time-invariant and included in the model as predictors for Wave 1 constructs only, covariates with less than marginally significant effects were excluded from the model one-by-one until only covariates with marginally significant effects ($p < .10$) remained (see Little, 2013). The final model gave a good model fit (χ^2 (315) = 556.271; RMSEA = .04; CFI = .97). The final covariate effects (excluded from Figure 5) included small effects of gender ($\beta = .10$) and mothers' education ($\beta = .15$) on school engagement, and a small effect of fathers' education ($\beta = .08$) on ISR. Furthermore, self-reported grades showed a strong effect ($\beta = .36$) on ISR and a very strong effect on school engagement ($\beta = .57$).

Figure 5 shows the final standardized structural model with cross-lagged paths constrained to be equal over time. The figure shows that school engagement and ISR were not stable over time. In fact, school engagement and ISR were significantly less stable during Grade 10 then during Grade 9 (Wald(2) = 41.99, *p* < .001). Significant cross-lagged effects (β = .09 to .21) were found for both school engagement and ISR across the three intervals observed in this study, confirming the hypothesized reciprocal effects between school engagement and ISR during adolescence. A follow-up test confirmed that the two groups of cross-lagged effects were not significantly different from each other (Wald(1) = 12.08, *p* = .72). However, as the effect size were small, and research has previously not been able to confirm both of the hypothesized cross-lagged effects in the same model, the two groups of cross-lagged effects were not constrained to be equal.

In sum, the final model, including three covariate variables, revealed significant cross-lagged effects across the three intervals for school engagement and ISR confirming the hypothesized reciprocal relation between the two constructs. Furthermore, the model-building process revealed a decrease in the stability of both school engagement and ISR from Grade 9 to Grade 10.
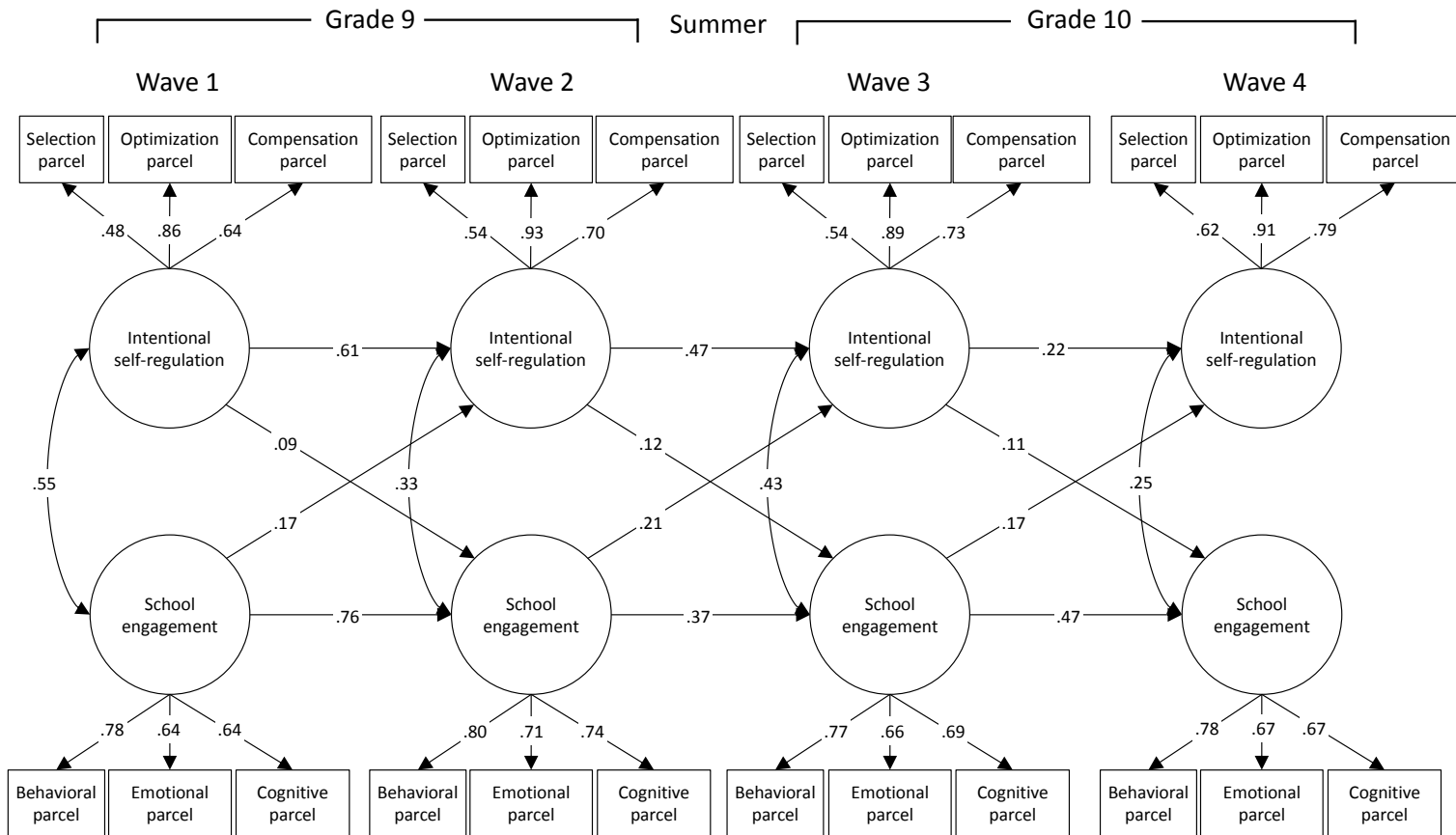
Figure 5. Cross-lagged panel model showing the standardized reciprocal influences between school engagement and ISR at the beginning and end of Grade 9 and Grade 10 in Iceland. The correlating residuals, covariates, and contextual effects are not shown to increase clarity. All the effects in the figure were significant at the *p* < .01 level.

4   Discussion

The overall aim of the doctoral project was to test the hypothesized positive reciprocal relation between school engagement and ISR during adolescence. This hypothesis was supported, thereby supporting theories that claim that ISR skills are a key element in the promotion of school engagement (Skinner & Pitzer, 2012). These findings are different from at least, two prior studies (Berger & Karabenick, 2011; Ning & Downing, 2010) that found relations between school engagement and ISR and tested, but did not confirm, the hypothesized positive reciprocal relation between school engagement and ISR.

The doctoral project described in this dissertation focused on (1) the operationalizing of ISR, (2) the operationalizing of school engagement, and (3) the relations between school engagement and ISR in adolescence. The first two aims were achieved by the two studies described in Papers I and II, and the third aim was achieved by the study described in Paper III, where a longitudinal reciprocal relation between school engagement and ISR was supported. The doctoral project as a whole, suggests that school engagement and ISR are mutually reinforcing during adolescence. In the following sections, the main results of each paper are discussed. A special emphasis is placed on Paper I as those results are currently only available in Icelandic.

4.1   Paper I: Addressing issues regarding the measurement of ISR

The main aim of the first paper was to investigate the factor structure and the convergent validity of a multidimensional measure of ISR skills called SOC. SOC involves three component processes called selection, optimization, and compensation (Freund & Baltes, 2002). Prior research from the U.S. has suggested that adolescents do not reliably differentiate between the three processes and proposed a way to improve the validity of the SOC measure in adolescence by modeling SOC as a nine-item single-factor instead of three six item factors (Gestsdottir et al., 2011; Gestsdottir & Lerner, 2007; Zimmerman et al., 2007). Prior to the current research a nine-item single-factor measure of SOC had not been supported with Icelandic 14 year old adolescents (Gestsdottir et al., 2011). As previous research has found the SOC measure to have problematic psychometric

qualities when used with U.S. and Icelandic youth (Geldhof, Bowers, et al., 2015; Geldhof, Gestsdottir, et al., 2015; Geldhof et al., 2012), various steps were taken to ensure the validity and reliability of the measure. The aim of the first paper was twofold. First, to investigate the benefits of using Likert-type answer options for the SOC measure instead of the original forced-choice answer options and, second, to conduct a confirmatory factor analysis (CFA) to test the presence of a nine-item single-factor of the SOC measure as proposed by earlier research.

The use of Likert-type answer options in measuring SOC had several benefits. The amount of error variance involved in the measure decreased (i.e., higher reliability) and the distribution of scores became more normal. A nine-item general factor structure was supported at both the beginning and the end of Grade 9, which indicated longitudinal configural invariance of the SOC measure during Grade 9. An earlier study on the forced-choice version of the SOC measure among 14-year-old students in Iceland did not confirm a nine-item general structure. Therefore, it seems that the adapted Likert-type answer options had a positive effect on the validity of the SOC measure compared to the earlier forced-choice version.

The results of Paper I indicated that Icelandic adolescents, like U.S. adolescents, did not reliably differentiate between the different subfactors of the SOC measure. In particular, Icelandic adolescents did not differentiate between items intended as manifestations of optimization and compensation, as the correlation between those subfactors was close to perfect ($r = .98$; see Paper I). The results of Paper I indicated that the factor structure of SOC existed as a general factor in adolescence but might be less developed during adolescence than in adulthood. Finally, the problems that adolescents had with answering negatively worded items during the pilot studies were pronounced, and might be indicative of a more general issue that young respondents have with negatively worded self-report questions. The use of negatively worded items with young respondents is a methodological concern that needs further research. The results of Paper I were important for the current study, as they informed a valid and reliable measure of ISR to test the main study hypothesis on the reciprocal relation between school engagement and ISR.

According to the findings presented in Paper I, adolescents did not reliably differentiate between the selection, optimization, and compensation subfactors of the SOC measure. The high correlation between the proposed subfactors might be an artifact of the wording of the questions. The questions were originally made for adults and might

therefore not be as well suited to reflect ISR in the lives of adolescents. Future research might consider taking a qualitative approach by observing, or interviewing adolescents (see e.g., Fredricks et al., 2016) in order to refine the conceptual understanding of ISR skills in adolescence.

4.2 Paper II: Addressing issues regarding the measurement of school engagement

The main aim of the second paper was to investigate the factor structure and the predictive validity of a multidimensional measure of school engagement. School engagement involves cognitive, emotional, and behavioral components that overlap conceptually. Prior research had suggested that, of the three components, only behavioral engagement predicted academic achievement more so than emotional engagement and cognitive engagement (see Chase et. al., 2014; Li & Lerner, 2010). However, the high correlation among the subfactors of school engagement indicated that the traditionally used correlated attributes model might be insufficient in explaining the variance observed in the scale items.

The study in Paper II revealed that academic achievement was best predicted by a general factor of school engagement that was manifested in all the school engagement items, regardless of their content origin. Furthermore, the bifactor analysis revealed that a substantial part of the effect on academic achievement, which had previously been ascribed to school engagement, was due to a factor not related to general school engagement, namely specific behavioral engagement. In other words, as shown by the latent factor of specific behavioral engagement, it is possible to attend school and finish homework, without showing signs of general school engagement, and such behavior also predicts academic achievement above and beyond general school engagement. Further research is needed to develop and confirm the existence of the specific behavioral engagement factor and its implications for academic achievement. The finding that academic achievement is most accurately predicted by modeling school engagement as a general factor is in contrast with previous findings, which indicated that, of the three school engagement components, only behavioral engagement predicted academic achievement (see Chase et al., 2014; Li et al., 2010).

In sum, the results of Paper II are consistent with Christenson et al. (2012, pp. 816-817) definition of school engagement as consisting behavioral (including academic), cognitive, and affective subtypes without any one subtype outweighing the other two in the definition of general

school engagement. The significantly better fit of the bifactor model suggests that, rather than being unidimensional or multidimensional, school engagement is characterized by both a single and multiple dimensions. Furthermore, the results showed that important secondary dimensions are lost (i.e., specific behavioral engagement), when using nested models, such as a three-factor model, to predict academic achievement. The results of Paper II were important for the current study, as they informed a holistic and valid measure of school engagement to test the main study hypothesis on the reciprocal relation between school engagement and ISR.

Shortly after the publication of Paper II, Wang et al. (2016) published a similar study that also supported the bifactor structure of school engagement. This second study took a domain specific approach and measured school engagement in math and science in middle and high school in the United States. The findings suggested that school engagement was comprised of multiple related yet distinct measures that supported a bifactor structural model. As such, the findings of Wang et al. (2016) further suggest that a bifactor model is the appropriate way to represent an integrated and holistic measure of school engagement.

The analysis in Paper II suggests that school engagement is a general factor that is simultaneously manifested in behavioral, emotional, and cognitive items. It is important to clarify that defining school engagement as a general factor does not mean that every meaningful manifestation of school engagement has been discovered. On the contrary, research into different subdimensions of school engagement should be encouraged, as they are likely to help researchers and educators to better understand school engagement conceptually. A more nuanced conceptual understanding of the multiple related subdimensions of school engagement can provide directions for researchers and educators to develop ways to promote school engagement that fit students with different strengths in environments with different assets. An example of such recent line of research can be found both in the work on social engagement (Fredricks et al., 2016), situational engagement (Shernoff et al., 2016), and the possible dark side of engagement (Salmela-Aro, Moeller, Schneider, Spicer, & Lavonen, 2016).

The dark side of engagement refers to findings that suggest that above average levels of school engagement do not necessarily relate to a positive school experience. Recent work on such possible dark sides of engagement is of special relevance to the current research for methodological and theoretical reasons. Using a person-centered approach Tuominen-Soini and

Salmela-Aro (2014) found that almost one in every four Finnish students experienced above levels of both school engagement and exhaustion simultaneously. This subgroup was also likely to develop depressive symptoms during the course of the study. Salmela-Aro et al. (2016) also pointed out that despite being engaged in school this subgroup might need specific support to maintain their motivation and resist burnout. The identification of this subgroup directly relates to the main hypothesis of the current study as these students (who may first not be seen at a risk of burnout) might nonetheless need specific support to develop their ISR skills to maintain their school engagement.

Methodologically, finding a subgroup of simultaneously engaged and exhausted students is of interest as it points out the limitations of the variable-centered approach used in most studies of school engagement, including the current one. It may be that a subgroup with high levels of school engagement and low levels of ISR skills exists in our data, although the normative trend points to a positive reciprocal relation between the two constructs. Further, person-centered analysis on the development of school engagement and ISR is needed on the current dataset to avoid falsely inferring variable-oriented research results to the person-level.

4.3 Paper III: The reciprocal relation between school engagement and ISR

The promotion of school engagement is widely considered a central focus for school reform and the design of learning environments (Shernoff, 2013). This focus stems from growing evidence that links school engagement negatively with school dropout and positively with indicators of academic achievement and well-being (Christenson et. al., 2012). In addition, it has been suggested that ISR skills are a key element in the promotion of school engagement (Skinner & Pitzer, 2012). However, to my knowledge, little or no direct evidence exists about the hypothesized reciprocal relations between school engagement and ISR skills.

The principal aim of Paper III was to test the hypothesis that a positive reciprocal relation existed between school engagement and ISR during the last two years of compulsory school in Iceland. This hypothesis was supported by modeling a cross-lagged relation between school engagement and ISR during the three intervals studied (see Figure 3). The findings indicated ISR as a part of one of the feedback loops relating to school engagement as suggested by the model of motivational dynamics (see Figure 2; Skinner & Pitzer, 2012).

Knowledge about the relative importance of school engagement and ISR has implications for those interested in promoting academic success in adolescence. Supporting the hypothesized reciprocal relation of school engagement and ISR suggests that ISR skills may be an important element in the promotion of school engagement and academic success (Skinner & Pitzer, 2012). Consequently, the current findings support further research on the inclusion of ISR skills in the design of optimal learning environments for school engagement.

The cross-lagged panel model also revealed differences in the stability of ISR across school years, where, as with school engagement, the stability of ISR during Grade 9 was high and significantly more stable than ISR during Grade 10. The decreased stability of ISR and school engagement during Grade 10 is consistent with theories that present school engagement and ISR as malleable constructs that are open to contextual conditions.

4.4   Future research

As previously mentioned, scholars have called for a better understanding on the degree of overlap between school engagement and ISR (Boekaerts, 2016; Eccles, 2016). The current research has provided support for the reciprocal relation between school engagement and ISR, thereby indicating their unique contribution to each other's development. In other words, school engagement and ISR seem strongly related but distinct concepts. However, many questions remain unanswered about ISR and school engagement during adolescence.

The addition of several education related covariates (i.e., gender, socio-economic status, parents' education and occupation, and self-reported grades) to the cross-lagged panel model highlighted the strong effect of prior academic achievement on later school engagement ($\beta = 0.57$). The strong relation between academic achievement and school engagement, together with the high stability of school engagement in Grades 9 and 10, gives reason to conduct further studies on the relation between academic achievement and school engagement during and prior to Grades 9 and 10. The model of motivation dynamics (see Figure 2) suggests a reciprocal relation between action (i.e., school engagement and ISR) and outcomes (i.e., academic achievement).

Prior research on the reciprocal relation between school engagement and academic achievement has partly supported this relation (see Chase et al., 2014) and prior research including ISR skills, school engagement, and academic achievement has pointed towards school engagement fully

mediating the effects of ISR skills on academic achievement (Li et al., 2010). However, both of these prior studies used the correlated attributes model in their analysis thereby not separating specific behavioral engagement from general school engagement (see Paper II). As demonstrated in Paper II, specific behavioral engagement predicted academic achievement beyond general school engagement. The use of the correlated attributes model in prior research including school engagement and academic achievement might therefore not give valid results as there is no way to account for the effects of the specific behavioral engagement variance using the correlated attributes model. Replicating the findings of Chase et al. (2014) and Li et al. (2010) using a bifactor model of school engagement would provide a more valid representation of school engagement and therefore be useful for further studies aimed at raising both school engagement and academic achievement.

Furthermore, research is needed to examine the possible existence of subgroups of students that do not fit the normative description of the variable-oriented analysis applied in the current study. This need could be fulfilled by identifying longitudinal subgroups using mixture modeling as described by Muthén and Muthén (1998-2012) and comparing the ISR subgroup composition of different school engagement subgroups and vice versa. Such a comparison would provide information about the extent to which the reciprocal relation at the between individual level established in the current research would apply to the within individual level.

School engagement has been considered malleable and subject to influences from teachers, parents, and peers (Skinner & Pitzer, 2012). However, school engagement has mostly been studied with stable measures that focus on the difference between individuals and treats moment-to-moment fluctuations in school engagement as measurement error (Salmela-Aro et al., 2016). As a teacher, I am very interested in why student engagement fluctuates within individuals from one learning situation to another as such information can be used to apply informed changes to the learning environment. Such research needs repeated measurement of situational school engagement to capture changes from one context to another. The experience sampling method (ESM; Hektner, Schmidt, & Csikszentmihalyi, 2007) is one method that future research could employ to capture the within individual differences in school engagement. The ESM involves the use smartphones or smartwatches, which are used to present short context specific questionnaires to the same individual. The questionnaires are presented many times during the day, for several days, to capture the effects of changing contexts.

As reviewed in the beginning of this synopsis the importance of contextual determinants for the development of school engagement has been well supported by empirical research. Prior research has, for example, identified teachers as an important facilitator of school engagement (Fredricks et al., 2004; Wang & Eccles, 2013). Teachers can provide warmth, structure and autonomy support by presenting relevant topics is an interesting way, setting challenging goals without overburdening students, and by providing formative feedback. Going back to the model of motivational dynamics (see Figure 2) it would be interesting to explore whether the warmth, structure, and autonomy support of teachers predicts the growth of both school engagement and ISR in similar ways. An imbalance in the effects of different types of support on school engagement and ISR might lead to the positive development of only one construct and not the other thereby reducing the potential to promote a positive integrated trajectory of school engagement and ISR. Such research could be implemented by using the previously described ESM and by applying person-oriented analysis such as the previously described mixture modeling.

4.5 Strengths and limitations

The study described in this synopsis had several strengths and limitations related to design, sampling, measures, and analyses. The main strength of the study is the longitudinal design that made it possible to predict dependent variables instead of only establishing relations, as in cross-sectional designs. In addition, the random, and sufficiently large sample afforded small, yet important, effect sizes (see Paper III) to be reliably inferred to a larger population. Furthermore, the measures used in this research were carefully selected, translated, piloted, and pretested according to recent developments in measurement and theory. Finally, the latest developments in structural equation modeling (SEM) were used to address issues of modeling longitudinal data, measurement error, multidimensionality, and handling missing data.

The main limitation of the study was that most of the measures were based on self-report, and are thus vulnerable to social desirability biases. Furthermore, using one form of data collection may have biased the results, as the observed correlations between the different items may be due to common method variance rather than representing actual relations among underlying constructs. An exception to the self-report data-collection, and a strength of the study, was the inclusion of Icelandic national examination

(INE) scores, which provided a more valid and normally distributed measure of academic achievement than the frequently used self-reported grades. The rigorous developmental phase of the ISR measure, in addition to the confirmatory factor analysis of the school engagement and ISR measures (see Papers I and II), indicated construct, content, and criterion validity. These findings reduce concerns related to using self-reported measures.

Another limitation of the current study is that all the analyses were variable-oriented. Such methods are dominant in psychological research and describe normative relations between variables. Variable-oriented research can provide important directions for future research but does not take into account the possibility of subgroups that might deviate from the normative trend. More person-oriented research is needed to support the current findings.

5   Conclusions

One of the reason many teachers, like myself, are interested in school engagement and ISR skills is because these concepts are within the sphere of our influence, or in other words, malleable (Appleton et al., 2008; Fredricks et al., 2004). Another reason is that these malleable concepts have been found to predict positive outcomes such as academic achievement (see e.g., Christenson et al., 2012; Tuominen-Soini & Salmela-Aro, 2014). However, research of school engagement and ISR have suffered from conceptual confusion and measurement issues and, as such, scholars have called for a better understanding on the degree of overlap between school engagement and ISR skills (Boekaerts, 2016; Eccles, 2016).

The current research has provided empirical findings to help clarify some of these issues. First, by assessing a modified, age-appropriate version of the SOC measure of ISR. Second, by providing empirical support for modeling school engagement as a bifactor model, thereby supporting theories that represent school engagement as an integrated and holistic concept. And third, by providing support for school engagement and ISR skills being highly related, yet distinct concepts, that mutually reinforce each other during adolescence.

In the current study, I have reviewed several theories from developmental and educational motivational psychology. Although the empirical findings of the current study only support limited parts of the theories reviewed, the review has made me think about the directions these theories provide for teaching and learning. Mosher and McGowan (1985) pointed out that students can be forced to attend school, but not forced to be engaged in school. The theories reviewed in the current research offer some direction on how teachers may be able to promote school engagement. The 5Cs model of PYD would suggest that teachers should make school meaningful to every student by aligning individual and contextual strengths to promote positive outcomes. The theory of motivational dynamics would suggest that the teachers needed to show warmth, and care about their students to promote relatedness. The same theory would suggest that teachers should emphasize competence by presenting relevant topics, setting challenging goals without overburdening students, and by providing formative feedback. Furthermore, the theory of motivational dynamics would suggest that teachers should support

autonomy by providing students with opportunities for voluntary choice, participation in school policy, and cooperation. As reviewed earlier in this synopsis these recommendations have empirical support although more research is needed. Finally, all the theoretical approaches reviewed in this doctoral project (i.e., the relational developmental systems metatheory, the PYD model, and the model of motivational dynamics) would emphasize the importance of supporting students' own activities and actions for increasing the probability of positive change. The high correlation and reciprocal relations between school engagement and ISR, found in the current research, support claims that ISR skills are one of the actions that are promoted by and promote school engagement. These findings support further research on the inclusion of ISR skills in the design of optimal learning environments for school engagement.

References

Ackerman, T. A. (1992). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. *Journal of Educational Measurement, 29*(1), 67-91. doi:10.1111/j.1745-3984.1992.tb00368.x

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools, 45*(5), 369-386. doi:10.1002/Pits.20303

Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist, 50*(1), 84-94. doi:10.1080/00461520.2015.1004069

Baltes, P. B. (1997). On the incomplete architecture of human ontogeny - Selection, optimization, and compensation as foundation of developmental theory. *American Psychologist, 52*(4), 366-380. doi:10.1037//0003-066X.52.4.366

Baltes, P. B., Baltes, M., Freund, A., & Lang, F. (1999). *The measurement of selection, optimization, and compensation (SOC) by self report. Technical report.* (978-3-87985-075-4). Retrieved from Berlin, Germany: http://tinyurl.sfx.mpg.de/sur8

Baltes, P. B., & Baltes, M. M. (1990). *Successful aging: Perspectives from the behavioral sciences*. NY: Cambridge University Press.

Bandura, A. (1986). *Social foundations of thought and action: a social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Adolescence and education* (pp. 306-337). Greenwich, CT: IAP - Information Age.

Berger, J.-L., & Karabenick, S. A. (2011). Motivation and students' use of learning strategies: Evidence of unidirectional effects in mathematics classrooms. *Learning and Instruction, 21*(3), 416-428. doi:10.1016/j.learninstruc.2010.06.002

Betts, J. (2012). Issues and methods in the measurement of student engagement: Advancing the construct through statistical modeling. In S. L. Christenson, L. A. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 783-805). NY: Springer.

Blondal, K., Jónasson, J., & Tannhäuser, A.-C. (2011). Dropout in a small society: Is the Icelandic case somehow different? In S. Lamb, E. Markussen, R. Teese, J. Polesel, & N. Sandberg (Eds.), *School Dropout and Completion* (pp. 233-251). Dordrecht: Springer.

Boekaerts, M. (2016). Engagement as an inherent aspect of the learning process. *Learning and Instruction, 43*, 76-83. doi:10.1016/j.learninstruc.2016.02.001

Bowers, E. P., Gestsdottir, S., Geldhof, G. J., Nikitin, J., von Eye, A., & Lerner, R. M. (2011). Developmental trajectories of intentional self regulation in adolescence: The role of parenting and implications for positive and problematic outcomes among diverse youth. *Journal of Adolescence, 34*(6), 1193-1206. doi:10.1016/j.adolescence.2011.07.006

Bowers, E. P., Wang, J., Tirrell, J. M., & Lerner, R. M. (2016). A cross-lagged model of the development of mentor-mentee relationships and intentional self-regulation in adolescence. *Journal of Community Psychology, 44*(1), 118-138. doi:10.1002/jcop.21746

Brandtstädter, J. (1998). Action perspectives on human development. In W. Damon & R. M. Richard (Eds.), *Handbook of child psychology: Theoretical models of human development* (5th ed., Vol. 1, pp. 807-863). New York, NY: Wiley.

Brandtstädter, J. (2006). Action perspectives on human development. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development* (6. ed., Vol. 1, pp. 516-568). New York, NY: Wiley.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. New York, NY: Guilford.

Chase, P. A., Hilliard, L. J., Geldhof, G. J., Warren, D. J. A., & Lerner, R. M. (2014). Academic achievement in the high school years: The changing role of school engagement. *Journal of Youth and Adolescence, 43*(6), 884-896. doi:10.1007/s10964-013-0085-4

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255. doi:10.1207/S15328007sem0902_5

Christenson, S. L., Reschly, L. A., & Wylie, C. (Eds.). (2012). *Handbook of research on student engagement*. New York, NY: Springer.

Cleary, T. J., & Zimmerman, B. J. (2012). A cyclical self-regulatory account of student engagement: Theoretical foundations and applications. In S. L. Christenson, L. A. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement*. NY: Springer

Coalition for Psychology in Schools and Education. (2006). *Report on the teacher needs survey*. American Psychological Association, Center for Psychology in Schools and Education. Washington, D.C.

Eccles, J. S. (2016). Engagement: Where to next? *Learning and Instruction, 43*, 71-75. doi:10.1016/j.learninstruc.2016.02.003

Eccles, J. S., & Gootman, J. A. (2002). *Community programs to promote youth development*. Washington, DC: National Academy Press.

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Maciver, D. (1993). Development during adolescence - The impact of stage-environment fit on young adolescents experiences in schools and in families. *American Psychologist, 48*(2), 90-101. doi:10.1037//0003-066x.48.2.90

Engels, M. C., Colpin, H., Van Leeuwen, K., Bijttebier, P., Van Den Noortgate, W., Claes, S., . . . Verschueren, K. (2016). Behavioral engagement, peer status, and teacher-student relationships in adolescence: A longitudinal study on reciprocal influences. *Journal of Youth and Adolescence, 45*(6), 1192-1207. doi:10.1007/s10964-016-0414-5

Farley, J. P., & Kim-Spoon, J. (2014). The development of adolescent self-regulation: Reviewing the role of parent, peer, friend, and romantic relationships. *Journal of Adolescence, 37*(4), 433-440. doi:10.1016/j.adolescence.2014.03.009

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109.

Fredricks, J. A., Filsecker, M., & Lawson, M. A. (2016). Student engagement, context, and adjustment: Addressing definitional, measurement, and

methodological issues. *Learning and Instruction, 43*, 1-4. doi:10.1016/j.learninstruc.2016.02.002

Freund, A. M., & Baltes, P. B. (1998). Selection, optimization, and compensation as strategies of life management: Correlations with subjective indicators of successful aging. *Psychology and Aging, 13*(4), 531-543.

Freund, A. M., & Baltes, P. B. (2002). Life-management strategies of selection, optimization, and compensation: Measurement by self-report and construct validity. *Journal of Personality and Social Psychology, 82*(4), 642-662. doi:10.1037//0022-3514.82.4.642

Gallup Student Poll. (2015). *Engaged today — ready for tomorrow*. Retrieved from http://www.gallupstudentpoll.com/

Geldhof, G. J., Bowers, E. P., Gestsdottir, S., Napolitano, C. M., & Lerner, R. M. (2015). Self-egulation across adolescence: Exploring the structure of selection, optimization, and compensation. *Journal of Research on Adolescence, 25*(2), 214-228. doi:10.1111/jora.12131

Geldhof, G. J., Gestsdottir, S., Stefansson, K. K., Johnson, S. K., Bowers, E. P., & Lerner, R. M. (2015). Selection, optimization, and compensation: The structure, reliability, and validity of forced-choice versus Likert-type measures in a sample of late adolescents. *International Journal of Behavioral Development, 39*(2), 171-185. doi:10.1177/0165025414560447

Geldhof, G. J., Little, T. D., & Hawley, P. H. (2012). Two measures of self-regulation for young adults and late adolescents in the academic and social domains. *International Journal of Behavioral Development, 36*(6), 476-488. doi:10.1177/0165025412462153

Gestsdottir, S., Adalbjarnardottir, S., & Thorsdottir, F. (2011). Formgerð sjálfsstjórnar: rannsókn meðal íslenskra ungmenna. *Tímarit sálfræðingafélags Íslands, 16*, 7-21.

Gestsdottir, S., Bowers, E., von Eye, A., Napolitano, C. M., & Lerner, R. M. (2010). Intentional self regulation in middle adolescence: The emerging role of loss-based selection in positive youth development. *Journal of Youth and Adolescence, 39*(7), 764-782. doi:10.1007/s10964-010-9537-2

Gestsdottir, S., Geldhof, G. J., Paus, T., Freund, A. M., Adalbjarnardottir, S., Lerner, J. V., & Lerner, R. M. (2015). Self-regulation among youth in four

Western cultures: Is there an adolescence-specific structure of the Selection-Optimization-Compensation (SOC) model? *International Journal of Behavioral Development, 39*(4), 346-358. doi:10.1177/0165025414542712

Gestsdottir, S., & Lerner, R. M. (2007). Intentional self-regulation and positive youth development in early adolescence: Findings from the 4-H study of positive youth development. *Developmental Psychology, 43*(2), 508-521. doi:10.1037/0012-1649.43.2.508

Gestsdottir, S., & Lerner, R. M. (2008). Positive development in adolescence: The development and role of intentional self-regulation. *Human Development, 51*(3), 202-224. doi:10.1159/000135757

Gestsdottir, S., Lewin-Bizan, S., von Eye, A., Lerner, J. V., & Lerner, R. M. (2009). The structure and function of selection, optimization, and compensation in middle adolescence: Theoretical and applied implications. *Journal of Applied Developmental Psychology, 30*(5), 585-600. doi:10.1016/j.appdev.2009.07.001

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology, 93*(1), 3-13. doi:10.1037//0022-0663.93.1.3

Heckhausen, J., & Tomasik, M. J. (2002). Get an apprenticeship before school is out: How German adolescents adjust vocational aspirations when getting close to a developmental deadline. *Journal of Vocational Behavior, 60*(2), 199-219. doi:10.1006/jvbe.2001.1864

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Thousand Oaks, CA: Sage.

Holsen, I., Geldhof, G. J., Larsen, T., & Aardal, E. (in press). The five Cs of positive youth development in Norway: Assessment and associations with positive and negative outcomes. *International Journal of Behavioral Development*, 1-11. doi:10.177/0165025416645668

IBM Corp. (2011). *IBM SPSS Statistics for Windows* (20.0 ed.). Armonk, NY: IBM Corp.

Icelandic Educational Testing Institute. (2014). The website of the Icelandic Educational Testing Institute. Retrieved from http://www.namsmat.is

Jang, H., Kim, E. J., & Reeve, J. (2016). Why students become more engaged or more disengaged during the semester: A self-determination theory dual-process model. *Learning and Instruction, 43*, 27-38. doi:10.1016/j.learninstruc.2016.01.002

Jarvela, S., Jarvenoja, H., Malmberg, J., Isohatala, J., & Sobocinski, M. (2016). How do types of interaction and phases of self-regulated learning set a stage for collaborative engagement? *Learning and Instruction, 43*, 39-51. doi:10.1016/j.learninstruc.2016.01.005

Karabenick, S., & Zusho, A. (2015). Examining approaches to research on self-regulated learning: conceptual and methodological considerations. *Metacognition and Learning, 10*(1), 151-163. doi:10.1007/s11409-015-9137-3

Kline, P. (1999). *The handbook of psychological testing*. London: Routledge.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3. ed.). New York, NY: Guilford Press.

Leaper, C., Farkas, T., & Brown, C. S. (2012). Adolescent girls' experiences and gender-related beliefs in relation to their motivation in Math/Science and English. *Journal of Youth and Adolescence, 41*(3), 268-282. doi:10.1007/s10964-011-9693-z

Lenth, R. V. (2001). Some practical guidelines for effective sample size determination. *American Statistician, 55*(3), 187-193.

Lerner, J. V., Bowers, E. P., Minor, K., Boyd, M. J., Mueller, M. K., Schmid, K. L., . . . Lerner, R. M. (2013). Positive youth development: processes, philosophies, and programs. In R. M. Lerner (Ed.), *Handbook of Psychology* (2nd ed., Vol. 6, pp. 752). Hoboken, NJ: Wiley.

Lerner, J. V., Phelps, E., Forman, Y., & Bowers, E. P. (2009). Positive youth development. In R. M. Lerner & L. D. Steinberg (Eds.), *Handbook of Adolescent Psychology* (3. ed., Vol. 1, pp. 524-558). Hoboken, NJ: Wiley.

Lerner, R. M. (2005). *Promoting positive youth development: Theoretical and empirical bases (White paper prepared for Workshop on the Science of Adolescent Health and Development).* National Research Council/Institute of Medicine. Washington, D.C.

Lerner, R. M., Lerner, J. V., Almerigi, J. B., Theokas, C., Phelps, E., Gestsdottir, S., . . . von Eye, A. (2005). Positive youth development, participation in community youth development programs, and community contributions of fifth-grade adolescents: Findings from the

first wave of the 4-H study of positive youth development. *Journal of Early Adolescence, 25*(1), 17-71. doi:10.1177/0272431604272461

Lerner, R. M., Lerner, J. V., & Benson, J. B. (2011). Positive youth development: Research and applications for promoting thriving in adolescence. *Advances in Child Development and Behavior, 41*, 1-17.

Lerner, R. M., Lerner, J. V., Bowers, E. P., & Geldhof, J. (2015). Positive youth development and relational-developmental-systems. In R. M. Lerner, M. H. Bornstein, & T. Leventhal (Eds.), *Handbook of Child Psychology and Developmental Science* (7th ed., Vol. 4). Hoboken, NJ: John Wiley & Sons, Inc.

Lerner, R. M., Lerner, J. V., von Eye, A., Bowers, E. P., & Lewin-Bizan, S. (2011). Individual and contextual bases of thriving in adolescence: A view of the issues. *Journal of Adolescence, 34*(6), 1107-1114. doi:10.1016/j.adolescence.2011.08.001

Lerner, R. M., von Eye, A., Lerner, J. V., Lewin-Bizan, S., & Bowers, E. P. (2010). The meaning and measurement of thriving: A view of the issues. *Journal of Youth and Adolescence, 39*(7), 707-719. doi:10.1007/s10964-010-9531-8

Li, Y. B. (2011). School engagement: What it is and why it is important for positive youth development. In R. M. Lerner, J. V. Lerner, & J. B. Benson (Eds.), *Advances in Child Development and Behavior* (Vol. 41, pp. 131-160). San Diego, CA: Academic press.

Li, Y. B., Lerner, J., & Lerner, R. (2010). Personal and ecological assets and academic competence in early adolescence: The mediating role of school engagement. *Journal of Youth and Adolescence, 39*(7), 801-815. doi:10.1007/s10964-010-9535-4

Li, Y. B., & Lerner, R. M. (2011). Trajectories of school engagement during adolescence: implications for grades, depression, delinquency, and substance use. *Developmental Psychology, 47*(1), 233-247. doi:10.1037/A0021307

Li, Y. B., & Lerner, R. M. (2012). *Testing across-group and longitudinal measurement equivalence of a tripartite measure of adolescent school engagement*. American Institutes for Research. Washington, DC.

Li, Y. B., & Lerner, R. M. (2013). Interrelations of behavioral, emotional, and cognitive school engagement in high school students. *Journal of Youth and Adolescence, 42*(1), 20-32. doi:10.1007/s10964-012-9857-5

Libbey, H. P. (2004). Measuring student relationships to school: Attachment, bonding, connectedness, and engagement. *Journal of School Health, 74*(7), 274-283. doi:10.1111/j.1746-1561.2004.tb08284.x

Little, T. D. (2013). *Longitudinal structural equation modeling*. New York, NY: Guilford.

Martin, A. J. (2007). Examining a multidimensional model of student motivation and engagement using a construct validation approach. *British Journal of Educational Psychology, 77*, 413-440. doi:10.1348/000709906x118036

McClelland, M. M., John Geldhof, G., Cameron, C. E., & Wanless, S. B. (2015). Development and self-regulation. In R. M. Lerner, W. F. Overton, & P. C. M. Molenaar (Eds.), *Handbook of Child Psychology and Developmental Science* (7th ed., Vol. 1). Hoboken, NJ: John Wiley & Sons, Inc.

Mcdonald. (1999). *Test theory: A unified treatment.* NJ: Lawrence Erlbaum Associates.

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research and Perspectives, 2*(4), 201-218. doi:10.1207/s15366359mea0204_1

Mosher, R., & McGowan. (1985). *Assessing student engagement in secondary schools: Alternative conceptions, strategies of assessing, and instruments.* University of Wisconsin, Research and Development Center. (ERIC Document Reproduction Service No. ED 272812).

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus User´s Guide* (7 ed.). Los Angeles, CA: Muthén & Muthén.

Newman, R. S. (2002). What do I need to do to succeed…when I don't understand what I'm doing!?: Developmental influences on students' adaptive help seeking. In J. Eccles & A. Wigfield (Eds.), *Development of Achievement Motivation* (pp. 285-306). San Diego, CA: Academic Press.

Ning, H. K., & Downing, K. (2010). The reciprocal relationship between motivation and self-regulation: A longitudinal study on academic performance. *Learning and Individual Differences, 20*(6), 682-686. doi:10.1016/j.lindif.2010.09.010

OECD. (2012a). *Education at a Glance 2012: OECD Indicators*: OECD Publishing.

OECD. (2012b). *PISA 2009 Technical Report*: OECD Publishing.

Overton, W. F. (2010). Life-span development: Concepts and issues. In R. M. Lerner (Ed.), *The handbook of life-span development* (Vol. 1, pp. 1-30). New Jersey, NJ: John Wiley & Sons, Inc.

Overton, W. F. (2013). A new paradigm for developmental science: relationism and relational-developmental systems. *Applied Developmental Science, 17*(2), 94-107. doi:10.1080/10888691.2013.778717

Overton, W. F. (2015). Processes, relations, and relational-developmental-systems. In R. M. Lerner, W. F. Overton, & P. C. M. Molenaar (Eds.), *Handbook of Child Psychology and Developmental Science* (7th ed., Vol. 1, pp. 9-63). Hoboken, NJ: Wiley.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology, 95*(4), 667-686. doi:10.1037/0022-0663.95.4.667

Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and Emotion, 28*(2), 147-169. doi:10.1023/B:MOEM.0000032312.95499.6f

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*(5), 667-696. doi:10.1080/00273171.2012.715555

Reschly, A. L., & Christenson, S. L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In S. L. Christenson, L. A. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 3-19). New York, NY: Springer.

Rose, T. (2015). *The end of average: How we succeed in a world that values sameness* (1st ed.). New York: HarperOne.

Roth, J. L., & Brooks-Gunn, J. (2003). What exactly is a youth development program? Answers from research and practice. *Applied Developmental Science, 7*(2), 94-111.

Rutter, M. (2006). Implications of resilience concepts for scientific understanding. *Resilience in Children, 1094*, 1-12. doi:10.1196/annals.1376.002

Salmela-Aro, K., Moeller, J., Schneider, B., Spicer, J., & Lavonen, J. (2016). Integrating the light and dark sides of student engagement using person-oriented and situation-specific approaches. *Learning and Instruction, 43*, 61-70. doi:10.1016/j.learninstruc.2016.01.001

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology - An introduction. *American Psychologist, 55*(1), 5-14. doi:10.1037/0003-066x.56.1.89

Shernoff, D. J. (2013). *Optimal learning environments to promote student engagement* (R. Levesque Ed. 1 ed.). New York, NY: Springer.

Shernoff, D. J., Kelly, S., Tonks, S. M., Anderson, B., Cavanagh, R. F., Sinha, S., & Abdi, B. (2016). Student engagement as a function of environmental complexity in high school classrooms. *Learning and Instruction, 43*, 52-60. doi:10.1016/j.learninstruc.2015.12.003

Sigthorsson, R. (2008). *Mat í þágu náms eða nám í þágu mats: Samræmd próf, kennsluhugmyndir kennara, kennsla og nám í náttúrufræði og íslensku í fjórum íslenskum grunnskólum.* (Ph.D.), Iceland University of Education, Reykjavik.

Silbereisen, R. K., & Lerner, R. M. (2007). *Approaches to positive youth development.* London: Sage.

Simmons, R. G., Burgeson, R., Carltonford, S., & Blyth, D. A. (1987). The impact of cumulative change in early adolescence. *Child Development, 58*(5), 1220-1234. doi:10.1111/j.1467-8624.1987.tb01453.x

Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom - Reciprocal effects of teacher-behavior and student engagement across the school year. *Journal of Educational Psychology, 85*(4), 571-581. doi:10.1037/0022-0663.85.4.571

Skinner, E. A., & Pitzer, J. R. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 21-44). New York, NY: Springer.

Tuominen-Soini, H., & Salmela-Aro, K. (2014). Schoolwork engagement and burnout among finnish high school students and young adults: profiles, progressions, and educational outcomes. *Developmental Psychology, 50*(3), 649-662. doi:10.1037/a0033898

Usher, E. L., & Pajares, F. (2008). Self-efficacy for self-regulated learning - A validation study. *Educational and Psychological Measurement, 68*(3), 443-463. doi:10.1177/0013164407308475

Wang, M. T., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy-value perspective to understand individual and gender differences in STEM fields. *Developmental Review, 33*(4), 304-340. doi:10.1016/j.dr.2013.08.001

Wang, M. T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction, 28*, 12-23. doi:10.1016/j.learninstruc.2013.04.002

Wang, M. T., Fredricks, J. A., Ye, F. F., Hofkens, T. L., & Linn, J. S. (2016). The math and science engagement scales: Scale development, validation, and psychometric properties. *Learning and Instruction, 43*, 16-26. doi:10.1016/j.learninstruc.2016.01.008

Wen, M., Su, S. B., Li, X. M., & Lin, D. H. (2015). Positive youth development in rural China: The role of parental migration. *Social Science & Medicine, 132*, 261-269. doi:10.1016/j.socscimed.2014.07.051

Wolters, C. A. (2003). Regulation of motivation: Evaluating an underemphasized aspect of self-regulated learning. *Educational Psychologist, 38*(4), 189-205. doi:10.1207/S15326985ep3804_1

Wolters, C. A., & Taylor, D. J. (2012). A self-regulated learning perspective on student engagement. In S. L. Christenson, L. A. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 635-651). NY: Springer

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*(2), 64-70. doi:10.1207/s15430421tip4102_2

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166-183. doi:10.3102/0002831207312909

Zimmerman, S. M., Phelps, E., & Lerner, R. M. (2007). Intentional self-regulation in early adolescence: Asessing the structure of selection, optimization, and compensation processes. *International Journal of Developmental Science, 1*(3), 272-299. doi:10.3233/dev-2007-1310

Zimmerman, S. M., Phelps, E., & Lerner, R. M. (2008). Positive and negative developmental trajectories in us adolescents: Where the positive youth development perspective meets the deficit model. *Research in Human Development, 5*(3), 153-165. doi:10.1080/15427600802274001

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's alpha, Revelle's beta, and McDonald's (omega H): Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika, 70*(1), 123-133. doi:10.1007/s11336-003-0974-7

Appendix A

Self-efficacy for self-regulated learning (Bandura, 2006).

Please rate how certain you are that you can do each of the things described below:

1. Get myself to study when there are other interesting things to do
2. Always concentrate on school subjects during class
3. Take good notes during class instruction
4. Use the library to get information for class assignments
5. Plan my schoolwork for the day
6. Organize my schoolwork
7. Remember well information presented in class and textbooks

Answer options: *Cannot do, Can almost not do, Moderately can do, Can do, Highly certain can do.*

Icelandic version:

Trú á eigin vinnubrögð í námi (Bandura, 2006).

Spurt er hversu vel þú treystir þér til að geta eftirfarandi:

1. haldið mér að námi þegar eitthvað annað áhugavert er í boði
2. alltaf einbeitt mér að námsefninu í kennslustundum
3. tekið góðar glósur í kennslustundum
4. notað bókasafnið til að afla upplýsinga fyrir skólaverkefni
5. gert áætlun um hvað ég geri í skólanum í dag
6. skipulagt skólavinnu mína
7. fest mér í minni upplýsingar sem ég fæ í kennslustundum og úr námsbókum

Svarmöguleikar: *Get ekki, Get eiginlega ekki, Get nokkurn veginn, Get eiginlega alveg, Get alveg.*

Appendix B

General SOC questionnaire 24 item short version, target items only (Freund & Baltes, 2002).

Elective selection

1. I concentrate all my energy on a few things
2. I always focus on the one most important goal at a given time
3. When I think about what I want in life, I commit myself to one or two important goals
4. I always pursue goals one after the other
5. When I decide upon a goal, I stick to it
6. I consider exactly what is important for me

Loss-based selection

1. When I can't do something as well as I used to, I think about what exactly is important to me
2. If I can't do something as well as before, I concentrate only on essentials
3. When I can't carry on as I used to, I direct my attention to my most important goal
4. When things don't work so well, I pursue my most important goal first
5. When I am not able to achieve something anymore, I direct my efforts at what is still possible
6. When I can no longer do something in my usual way, I think about what, exactly, I am able to do under the circumstances

Optimization

1. I keep working on what I have planned until I succeed
2. If something matters to me, I devote myself fully and completely to it
3. I do everything I can to realize my plans
4. When I choose a goal, I am also willing to invest much effort in it

5. When something is important to me, I don't let setbacks discourage me

6. I think about when exactly I can best realize my plans

Compensation

1. For important things, I pay attention to whether I need to devote more time or effort

2. When things aren't going so well, I accept help from others

3. When things don't work the way they used to, I look for other ways to achieve them

4. When I can't do something as well as I used to, then I ask someone else to do it for me

5. When something doesn't work as well as usual, I look at how others do it

6. When something does not work as well as before, I listen to advisory broadcasts and books as well

Answer options: *Just like me, Somewhat like me, Neither like me or not like me, Not like me, and Not at all like me.*

Icelandic version:

Almenni SOC spurningalistinn 24 atriði, stutt útgáfa, einungis atriði sem lýsa sjálfstjórnun (Freund & Baltes, 2002).

Val byggt á missi

1. Þegar ég get ekki gert eitthvað eins vel og ég var vön/vanur, þá hugsa ég um hvað nákvæmlega skiptir mig miklu máli

2. Ef ég get ekki gert eitthvað eins vel og áður, þá einbeiti ég mér bara að meginatriðum

3. Þegar ég get ekki haldið áfram eins og ég var vön/vanur, þá beini ég athyglinni að mikilvægasta markmiðinu mínu

4. Þegar hlutirnir ganga ekki sem best, þá keppi ég að mikilvægasta markmiðinu mínu fyrst

5. Þegar ég get ekki lengur gert eitthvað þá einbeiti ég mér að því sem er enn mögulegt

6. Þegar ég get ekki lengur gert eitthvað eins og ég er vön/vanur, þá hugsa ég um hvað nákvæmlega ég er fær um að gera undir þessum kringumstæðum

Hámörkun

1. Ég geri allt hvað ég get til að ná markmiðum mínum

2. Til þess að ná markmiði mínu reyni ég eins margar nýjar leiðir eins og á þarf að halda

3. Þegar ég ætla mér að ná erfiðu markmiði bíð ég eftir réttu augnarbliki og besta tækifærinu

4. Þegar ég byrja á einhverju sem mér finnst skipta máli en veit að verður erfitt, þá legg ég sérstaklega hart að mér

5. Þegar ég vil ná árangri, þá skoða ég líka hvernig aðrir hafa gert það

6. Ég íhuga vandlega hvernig ég get best náð markmiðum mínum

Uppbót

1. Þegar eitthvað gengur ekki jafn vel og venjulega athuga ég hvernig aðrir hafa farið að

2. Þegar eitthvað virkar ekki eins vel og áður, fæ ég ráðleggingar á netinu eða les bók

3. Þegar mér finnst eitthvað vera mikilvægt þá velti ég fyrir mér hvort ég þurfi að verja meiri tíma í það eða að vera duglegri

4. Þegar hlutirnir ganga illa, þigg ég hjálp frá öðrum

5. Þegar hlutirnir ganga ekki jafn vel og áður reyni ég að gera þá öðruvísi

6. Þegar ég get ekki leyst eitthvað verkefni eins vel og áður þá bið ég einhvern annan að gera það fyrir mig

Svarmöguleikar: *Mjög líkt mér, Frekar líkt mér, Hvorki líkt né ólíkt mér, Frekar ólíkt mér, Mjög ólíkt mér.*

Appendix C

School engagement questionnaire 15 item (Li & Lerner, 2012).

Note:Items with * are reverse-coded.

Behavioral School Engagement (0 = *Never* to 3 = *Always*)

1. How often do you come to class unprepared (homework unfinished, forget to bring books or other materials, etc.)?*
2. How often do you complete homework on time?
3. How often do you skip classes without permission?*
4. How often do you actively take part in group (class) discussions?
5. How often do you work hard to do well in school?

Emotional School Engagement (0 = *Strongly disagree* to 3 = *Strongly agree*)

1. I feel part of my school
2. I care about the school I go to
3. I am happy to be at my school
4. I don't find school fun and exciting*
5. I enjoy the classes I am taking

Cognitive School Engagement (0 = *Strongly disagree* to 3 = *Strongly agree*)

1. I want to learn as much as I can at school
2. I think it is important to make good grades
3. I think the thing I learn at school are useful
4. I think a lot about how to do well in school
5. School is very important for later success

Icelandic version:

Ath: Atriði með * eru kóðuð öfugt.

Hegðunarleg virkni í skólastarfi (0 = *Aldrei* til 3 = *Oftast*)

1. Hversu oft kemur þú óundirbúinn í tíma (t.d. heimavinnu ólokið, gleymir bókum eða öðru efni)?*
2. Hversu oft klárar þú heimavinnu á réttum tíma?
3. Hversu oft sleppir þú tímum án leyfis (skrópar)?*

4. Hversu oft tekur þú virkan þátt í umræðum í hóp (bekk)?

5. Hversu oft leggur þú hart að þér til að standa þig vel í skólanum?

Tilfinningaleg virkni í skólastarfi (0 = *Mjög sammála* til 3 = *Mjög ósammála*)

1. Mér finnst ég hluti af skólanum mínum.

2. Mér er annt um skólann sem ég geng í.

3. Ég er ánægð/ur að vera í skólanum mínum.

4. Mér finnst skólinn ekki skemmtilegur eða spennandi.*

5. Ég hef gaman af fögunum sem ég er í.

Vitsmunaleg virkni í skólastarfi (0 = *Mjög sammála* til 3 = *Mjög ósammála*)

1. Ég vil læra eins mikið og ég get í skólanum.

2. Ég tel það mikilvægt að fá góðar einkunnir.

3. Ég tel það sem ég læri í skólanum vera gagnlegt.

4. Ég hugsa mikið um hvernig ég get staðið mig vel í skólanum.

5. Skólinn er mjög mikilvægur fyrir velgengni í framtíðinni.

Paper I: Kristján Ketill Stefánsson, Steinunn Gestsdóttir, & Sigurgrímur Skúlason. (2014). Þróun og mat á réttmæti mælitækis á meðvitaðri sjálfsstjórnun ungmenna. *Sálfræðiritið*, **19, 41-55.**

Þróun og mat á réttmæti mælitækis
á meðvitaðri sjálfstjórnun ungmenna

Kristján Ketill Stefánsson og Steinunn Gestsdóttir
Háskóli Íslands
Sigurgrímur Skúlason
Námsmatsstofnun

Meðvituð sjálfstjórnun vísar til getu fólks til að setja sér markmið og nota árangursríkar leiðir til að ná markmiðum sínum. Rannsóknir á slíkri getu benda til mikilvægis þess að hún sé til staðar á unglingsárum en ekki liggja fyrir nægilega réttmæt og áreiðanleg mælitæki sem henta unglingum. Markmið þessarar rannsóknar var tvíþætt. Í fyrsta lagi var athugað hvort hægt væri að bæta áreiðanleika sjálfstjórnunarmælitækisins SOC með því að setja svarmöguleika mælitækisins fram á Likert kvarða í stað tvíkosta svarmöguleika. Í öðru lagi var kannað hvort breytt framsetning mælitækisins yrði til þess að unnt væri að staðfesta formgerð þess og þar með hugtakaréttmæti við íslenskar aðstæður. Undirbúningsmæling með háskólanemum ásamt undirbúningsmælingu og rýnihópsviðtali við unglinga bentu til þess að Likert framsetning gæfi réttmætari niðurstöður en tvíkosta framsetning líkt og notuð hefur verið með fullorðnum þátttakendum. Staðfestandi þáttagreining á gögnum frá 539 nemendum í 9. bekk (46% stúlkur, meðalaldur 14,3 ár) leiddi í ljós að þríþátta formgerð SOC mælitækisins kom ekki fram. Einsþátta líkan féll aftur á móti vel að gögnunum líkt og sýnt hefur verið fram á með bandarískum ungmennum. Þessar niðurstöður benda til þess að einsþátta útgáfa SOC mælitækisins með Likert svarmöguleikum gefi áreiðanlega og réttmæta mælingu á sjálfstjórnun unglinga við upphaf og lok 9. bekkjar á Íslandi og gefi þar af leiðandi möguleika á samanburðarhæfum niðurstöðum við niðurstöður mælinga á bandarískum unglingum.

Efnisorð: Sjálfstjórnun, SOC, staðfestandi þáttagreining, mælitæki, unglingar.

Rannsóknir á meðvitaðri sjálfstjórnun (e. *intentional self-regulation*) hafa gefið til kynna sterk jákvæð tengsl við farsælan þroska og talið er að meðvituð sjálfstjórnun spili lykilhlutverk í þroskaferli unglinga (Lerner, Phelps, Forman og Bowers, 2009). Í þessari grein verður fjallað um tilurð, aðlögun og þróun mælitækis sem mælir meðvitaða sjálfstjórnun meðal ungs fólks á Íslandi. Markmið þessarar rannsóknar var tvíþætt. Í fyrsta lagi var athugað hvort hægt væri að bæta áreiðanleika og samleitniréttmæti sjálfstjórnunarmælitækisins SOC með því að setja svarmöguleika mælitækisins fram á Likert kvarða í stað tvíkosta framsetningar. Í öðru lagi var kannað hvort hægt væri að staðfesta eins eða þriggja þátta formgerð mælitækisins þegar notast var við Likert kvarða, og þar með lagt mat á hugtakaréttmæti þessarar útgáfu mælitækisins.

Meðvituð sjálfstjórnun

Sjálfstjórnun (e. *self-regulation*) vísar til hæfni fólks til að stjórna eigin hugsun, tilfinningum og hegðun og er talin mikilvægur þáttur í farsælli þroskaframvindu ungs fólks (Brandstädter, 2006). Samkvæmt kenningum um kerfi tengsla og þroska (e. *relational-developmental systems*) hafa einstaklingurinn og umhverfi hans gagnkvæm áhrif hvort á annað (Overton, 2013). Meðvituð (e. *intentional*) sjálfstjórnun hefur verið skilgreind sem þeir eiginleikar og þær leiðir sem fólk notar til að hafa áhrif á þetta samspil með því að setja sér þau markmið sem það vill ná, gera áætlanir um hvernig markmiðum verður náð og hvernig það stjórnar hugsun og hegðun til að ná markmiðum sínum (Gestsdottir og Lerner, 2008). Dæmi um meðvitaða sjálfstjórnun í námi er þegar nemendur finna upp eða nýta sér gagnlegar leiðir, til dæmis minnistækni, til að undirbúa sig undir próf (sem er markmið sem viðkomandi hefur sett sér). Sömu leiðirnar virka ekki endilega fyrir alla nemendur (ólíkir eiginleikar) eða frá einni skólastofu til annarrar (ólíkt umhverfi) en sérhver meðvituð leið í sérhverju samspili er dæmi um sjálfstjórnun að námsmarkmiði. Eins og sjá má gefur sjálfstjórnunarfærni fólki færi á að stjórna eigin þroskaferli og ná þeim markmiðum sem það telur muni bæta líf sitt (Baltes, 1997).

Sjálfstjórnunarfærni eykst með aldri í takt við þau markmið sem unglingar setja sér (Massey, Gebhardt og Garnefski, 2008). Þær breytingar sem verða á vitsmunum, líkams- þroska og umhverfi verða til þess að mikilvægi meðvitaðar sjálfstjórnunar eykst á unglingsárum og sjálfstjórnun verður mark- vissari, flóknari og tengist í auknum mæli langtímamarkmiðum (Brandstädter, 2006). Sem dæmi um breytingar má nefna auknar kröfur um sjálfstæð vinnubrögð og óhlutbundna hugsun eftir því sem líður á unglings- árin ásamt því að valnámskeið verða stærri hluti af námskránni (Mennta- og menningar- málaráðuneytið, 2011).

Sjálfstjórnun fólks á ólíkum aldri hefur á undanförnum árum vakið mikla athygli fræðimanna (McClelland, Ponitz, Messersmith og Tominey, 2010). Í menntarannsóknum er algengt að horft sé til breytinga á lærdóms-umhverfi nemenda á unglingastigi, svo sem breytinga á kennsluháttum eða breytinga á námsframboði, og hver áhrif þeirra á viðhorf eða námsárangur nemenda séu. Rannsóknir á hlutverki nemenda í eigin námsframvindu eru skemur á veg komnar. Geta nemenda til að hafa meðvituð áhrif á umhverfi sitt, og þar með þá stefnu sem þroski þeirra og nám tekur, er talin leika lykilhlutverk í skólagöngunni. Því er mikilvægt að auka skilning á þeirri færni sem nemendur beita til að stýra eigin námsárangri og hvernig styðja má við þá færni (McClelland og Cameron, 2011).

Rannsóknir á meðvitaðri sjálfstjórnun ungmenna

Rannsóknir hafa sýnt fram á mikilvægi sjálfstjórnunar fyrir námsárangur í leikskóla og á yngsta og miðstigi grunnskóla (McClelland, Acock og Morrison, 2006; McClelland, Morrison og Holmes, 2000). Algengt er að rannsóknir á sjálfstjórnun á yngri stigum fari fram með því að leggja mat á ákveðna hegðun. Dæmi um slíkt væri frestun á umbun *(e. delay of gratification)* sem talin er vera birtingarmynd *(e. proxy)* sjálfstjórnunar. Á unglingastigi er algengt að rannsóknir á sjálfstjórnun beinist að tengslum við langtímamarkmiðssetningu, áætlunargerð, gildismati, væntingum eða hugsanlegum afleiðingum (McClelland o.fl., 2010; Zimmerman, 2002).

Rannsóknir á sjálfstjórnun á unglingsárum eru skammt á veg komnar, meðal annars vegna þess að ekki hafa legið fyrir nægilega góð mælitæki á flókinni gerð sjálfstjórnunar fyrir þennan aldur (Geldhof, Bowers, Gestsdottir, Napolitano og Lerner, í prentun). Þær rannsóknir sem gerðar hafa verið hafa oftast beinst að stjórnun tilfinninga eða tengslum við áhættuhegðun (sjá til dæmis

Brody og Ge, 2001; Quinn og Fromme, 2010). Ein undantekning frá þessu er 4-H rannsóknin á farsælli þroskaframvindu (*4-H Study of Positive Youth Development*). 4-H rannsóknin er umfangsmikil rannsókn á þróun og hlutverki meðvitaðrar sjálfstjórnunar meðal unglinga. Markmið rannsóknarinnar var að meta hvaða eiginleikar og leiðir hjá ungmennum í sampili við hvaða umhverfi leiða til farsællar þroskaframvindu. Gagnasöfnun stóð yfir frá 2002-2010 og fylgdi eftir ríflega 7000 nemendum í 5.-12. bekk í 13 ríkjum Bandaríkjanna (Lerner, Phelps, Forman og Bowers, 2009).

Í 4-H rannsókninni var notast við mælitæki á meðvitaðri sjálfstjórnun sem kallast SOC (*Selection, Optimization, Compensation*). Það á uppruna sinn í rannsóknum Baltes og samstarfsmanna hans á sviði öldrunarfræða og hefur mikið verið notað með fullorðnum (Baltes, 1997; Freund og Baltes, 2002). SOC mælitækið mælir meðvitaða sjálfstjórnun sem býr að baki þeim markmiðum sem fólk setur sé í lífinu og þeim leiðum sem fólk fer til að ná þeim. Mælingin er fjórþætt og inniheldur spurningar um hversu vel fólki gengur að velja sér markmið (val; e. *selection*), hversu vel því gengur að finna leiðir að markmiðum sínum (hámörkun; e. *optimization*), hversu vel því gengur að skipta út leiðum ef þörf krefur (uppbót; e. *compensation*) og hversu vel því gengur að breyta markmiðum ef fyrri markmið reynast óraunhæf (val byggt á tapi; e. *loss based selection*). Sýnt hefur verið fram á viðunandi áreiðanleika mælitækisins í rannsóknum á fullorðnum í Þýskalandi (Freund og Baltes, 2002; Wiese, Freund og Baltes, 2000). Fjórða þættinum (val byggt á tapi) var sleppt meðal ungra svarenda í 4-H rannsókninni þar sem sá þáttur var ekki talinn eiga við unglingsárin sem einkennast af vexti á flestum sviðum þroska.

SOC mælingin sem notuð var í 4-H rannsókninni samanstóð af 18 tvíkosta atriðum (sex atriðum fyrir hvern undirþátt) þar sem svarendur þurftu að velja hvort þeir líktust frekar manneskju A eða B. Hvert atriði lýsti annars vegar hegðun sem er dæmigerð fyrir

sjálfstjórnun og hins vegar hegðun sem telst vera andstæða sjálfstjórnunar. Rannsóknir á gögnum úr 4-H rannsókninni sýndu að líkan sem gerði ráð fyrir öllum þremur þáttum SOC mælingarinnar féll illa að 4-H gögnunum meðal bandarískra unglinga. Þegar einsþátta líkan var metið sýndu niðurstöður að mátgæði þess voru viðunandi ef níu atriði voru fjarlægð úr mælitækinu. Þau níu atriði sem eftir stóðu (sjá 1. töflu) mynduðu eina almenna sjálfstjórnunarmælingu með lágum áreiðanleika í fimmta bekk ($\alpha = 0{,}55$) en betri áreiðanleika meðal unglinga í 7.-10. bekk ($\alpha = 0{,}62$-$0{,}65$) (Gestsdottir og Lerner, 2007; Gestsdottir, Lewin-Bizan, von Eye, Lerner og Lerner, 2009; Zimmerman, Phelps og Lerner, 2007). Þessar niðurstöður benda til þess að uppbygging SOC mælingarinnar sé ólík á unglings- og fullorðinsárum.

Niðurstöður 4-H rannsóknarinnar sýndu jákvæða fylgni milli SOC og einkenna farsæls þroska (sem dæmi um einkenni farsæls þroska má nefna sjálfstraust, umhyggju, siðferðisþrek og námsgetu) og neikvæða fylgni við neikvæða líðan og hegðun á unglingastigi (svo sem áfengisneyslu og þunglyndi, sjá nánar Gestsdottir og Lerner, 2007). Sem dæmi má nefna að í sjötta bekk reyndist fylgnin á milli SOC og getu í skóla vera jákvæð og miðlungs-sterk, $r(1312) = 0{,}34$, $p < 0{,}001$ (Bowers o.fl., 2011; Steinunn Gestsdóttir og Lerner, 2007).

SOC mæling úr 4-H rannsókninni var þýdd á íslensku og notuð í rannsókn með 14 og 18 ára unglingum á Íslandi árið 2009. Staðfestandi þáttagreining sýndi að þriggja þátta líkan féll illa að gögnunum en að sjö atriði (sjá 1. töflu) mynduðu eina sjálfstjórnunarmælingu með viðunandi mátgæðum sem er svipað og raunin var meðal bandarískra ungmenna (Steinunn Gestsdóttir, Sigrún Aðalbjarnardóttir og Fanney Þórsdóttir, 2011).

Aðferðafræðileg vandamál við mælingar á meðvitaðri sjálfstjórnun

Nokkur aðferðafræðileg vandamál hafa komið í ljós þegar SOC mælingin hefur verið notuð með ungmennum. Eins og bent hefur verið á hefur innri stöðugleika verið ábótavant og að auki hefur SOC mælingin ekki aðgreinst í þrjá undirþætti meðal unglinga. Jafnframt hefur áreiðanleiki einsþátta líkans ekki reynst góður, hvorki í mælingum með íslenskum né bandarískum unglingum. Nýleg leitandi þáttagreining sem byggðist á átta ára langtímamælingum úr áðurnefndu 4-H gagnasafni leiddi í ljós að til hliðar við hinn almenna SOC undirþátt myndaðist einnig aðferðafræðilegur undirþáttur fyrir þau atriði sem voru neikvætt orðuð í framsetningu mælitækisins (Geldhof o.fl., í prentun). Svipaðar niðurstöður fengust í nýlegri rannsókn með unglingum frá fjórum vestrænum löndum (Gestsdottir o.fl., 2014). Þetta gefur til kynna að orðalag sumra spurninganna reynist svarendum erfitt og að breyta þurfi orðalagi og/eða framsetningu spurninganna til að þær falli betur að ungum svarendum. Jafnframt hefur verið mælt með því að beita Likert svarmöguleikum við framsetningu SOC spurninganna til að aðgreina betur svarendur með litla og mikla sjálfstjórnun (Geldhof, Little og Hawley, 2012)

Þróun mælingar á meðvitaðri sjálfstjórnun fyrir unglinga

Rannsóknin sem hér er til umræðu er hluti af stærri rannsókn á þróun sjálfstjórnunar og tengslum hennar við farsælan þroska ungmenna á Íslandi. Rannsóknin hófst haustið 2011 og er stýrt af dr. Steinunni Gestsdóttur við Sálfræðideild Háskóla Íslands. Rannsóknin er langtímarannsókn þar sem 561 nemanda er fylgt eftir með spurningalista við upphaf og lok níunda og tíunda bekkjar.

Tilgangur þess hluta sem hér er lýst var að þróa mælingu á meðvitaðri sjálfstjórnun fyrir unglinga. Fyrri rannsóknir á sjálfstjórnun unglinga með notkun SOC mælitækisins hafa í flestum tilvikum einungis notað níu atriði (sjá 1. töflu) úr mælitækinu þar sem fjölþátta formgerð mælitækisins hefur ekki komið skýrt fram í mælingum á unglingum þrátt fyrir að hafa verið staðfest meðal fullorðinna. Erlendar rannsóknir hafa staðfest að einsþátta formgerð hentar unglingum frekar en fjölþátta formgerð (Gestsdottir o.fl., 2014; Gestsdottir o.fl., 2009; Zimmerman o.fl., 2007) en hvorug formgerðin kom fram í fyrri rannsókn með unglingum hér á landi (Steinunn Gestsdóttir o.fl., 2011).

Í þessari rannsókn voru þrjár útgáfur af SOC mælingum á meðvitaðri sjálfstjórnun prófaðar áður en endanleg útgáfa var valin til notkunar í rannsókninni. Einnig var skyld mæling á sjálfstjórnun í námi (Bandura, 2006) notuð til að leggja mat á samleitniréttmæti við SOC mælingarnar. Þróun og prófun á mælitækinu fór fram í fimm skrefum:

1. Fyrsta skrefið fólst í undirbúnings-mælingu meðal háskólanema í grunn-námi þar sem dreifing, áreiðanleiki og samleitniréttmæti niðurstaða úr Likert og tvíkosta framsetningum var borin saman.

2. Annað skrefið fólst í rýnihópsviðtali við hóp unglinga þar sem mat var lagt á framsetningu spurninga og svar-möguleika SOC mælitækisins.

3. Þriðja skrefið fólst í undirbúnings-mælingu II sem var framkvæmd með hópi unglinga til að skoða áreiðanleika, dreifingu og samleitniréttmæti SOC mælingarinnar með breyttum Likert svarmöguleikum.

4. Í fjórða skrefinu var gerð forprófun á lokaútgáfu spurningalistans með nýjum hópi unglinga til að staðfesta yfir-borðsréttmæti mælitækisins sem hluta af stærri spurningalista.

5. Fimmta skrefið fólst í staðfestandi þáttagreiningu á SOC mælingunni sem byggir á fyrstu tveimur gagnasöfnunum langtímarannsóknarinnar sem rannsóknin er hluti af. Þetta síðasta skref var framkvæmt til að staðfesta hugtakaréttmæti og langtímastöðugleika mælitækisins.

Skortur á áreiðanlegu og réttmætu mælitæki fyrir sjálfstjórnun meðal unglinga á Íslandi var kveikjan að eftirfarandi rannsóknarspurningum.

1. Er hægt að auka áreiðanleika og réttmæti SOC sjálfstjórnunarmælingarinnar meðal ungmenna á Íslandi með því að nota einungis þann hluta atriðanna sem lýsa sjálfstjórnun, sleppa þeim atriðum sem lýsa andstæðu sjálfstjórnunar og setja svarmöguleika mælingarinnar fram á Likert kvarða líkt og fyrri rannsóknir hafa kallað eftir (Geldhof o.fl., 2012)?

2. Ef breytt framsetning SOC mælitækisins leiðir til áreiðanlegri niðurstaða en fyrri mælingar hafa gefið til kynna er þá mögulegt að staðfesta eins- eða fjölþátta formgerð mælitækisins við íslenskar aðstæður?

Aðferð

Söfnun gagnanna fór fram í nokkrum skrefum svo unnt væri að þróa SOC mælitækið og um leið svara þeim rannsóknarspurningum sem settar voru fram. Gerð er grein fyrir aðferðafræði rannsóknarinnar í sömu röð og söfnun gagnanna fór fram.

Þátttakendur

Undirbúningsmæling I. Í undirbúningsmælingu I var notað hentugleikaúrtak þar sem háskólanemar í tveimur aðferðafræðinámskeiðum voru beðnir að svara rafrænum spurningalista. Að loknum þremur áminningum í tölvupósti höfðu 139 nemendur (92%

kvenkyns, meðalaldur 29,6 ár, staðalfrávik 8,2 ár) svarað könnuninni. Svarhlutfall var 32%.

Rýnihópsviðtal. Þátttakendur í rýnihópsviðtalinu voru 15 unglingar (33% stúlkur) í 9. bekk í skóla á Reykjavíkursvæðinu. Þátttakendurnir voru á 14. aldursári.

Undirbúningsmæling II. Þátttakendur í undirbúningsmælingu II voru 42 unglingar (50% stúlkur) í 9. bekk í skóla á Reykjavíkursvæðinu. Þátttakendurnir voru á 14. aldursári.

Forprófun. Þátttakendur í forprófun spurningalistans í heild sinni voru 77 unglingar (47% stúlkur) í 9. bekk í skóla á Reykjavíkursvæðinu. Þátttakendurnir voru á 14. aldursári.

Þátttakendur í langtímarannsókn. Þýði rannsóknarinnar (N = 2836) var takmarkað við meðalstóra og stóra skóla (með meira en 20 nemendur í 9. bekk) á höfuðborgarsvæðinu og Reykjanesi. Þýðið var takmarkað til að auðvelda framkvæmd rannsóknarinnar og lágmarka langtímabrottfall. Þegar rannsóknin hófst innihélt þýðið tvo þriðju hluta allra nemenda í 9. bekk á landinu. Tuttugu af 54 mögulegum skólum voru valdir af handahófi. Fimmtán skólar af tuttugu samþykktu að taka þátt í rannsókninni (75%). Tveir 9. bekkir voru valdir af handahófi innan hvers skóla til að tryggja góða dreifingu á milli skóla. Alls voru 625 nemendur skráðir í þá 30 bekki sem þátt tóku í rannsókninni. Í heildina gaf 561 foreldri skriflegt samþykki (90%) fyrir þátttöku barns síns. Svarhlutfall að lokinni fyrstu gagnasöfnun var 96% (N = 539, 46% stúlkur, meðalaldur 14,3 ár, staðalfrávik 102 dagar) og 93% (N = 519, 46% stúlkur, meðalaldur 14,8 ár) að lokinni annarri gagnasöfnun.

Framkvæmd

Undirbúningsmæling I. Leyfi var fengið hjá umsjónarmanni tveggja aðferðafræði-námskeiða til að leggja undirbúningsmælingu I fyrir þátttakendur námskeiðanna í byrjun maí 2012. Umsjónarmaður námskeiðanna lét nemendur sína vita um fyrirhugaða könnun og tilgang hennar og hvatti nemendur til að svara könnuninni. Könnunin var sett upp í K2 kannanakerfinu á innri vef Háskóla Íslands. Leitað var eftir athugasemdum um þau mælitæki sem prófuð voru með því að koma fyrir opnum textareit aftast í spurningalistanum með spurningunni „Ef þú vilt koma einhverju á framfæri varðandi mælitækin sem verið er að forprófa, vinsamlegast skrifaðu það í reitinn hér að neðan." Engum persónuupplýsingum var safnað við fyrirlögnina. Þrjár tölvupósts-áminningar voru sendar með reglulegu millibili til þeirra sem ekki voru búnir að svara á tveggja vikna tímabili.

Rýnihópsviðtal. Leyfi var fengið hjá skólastjórnendum í skóla í Reykjavík til að fá nemendur til að svara og ræða um leið í sameiningu um spurningalista rann-sóknarinnar. Viðtalið fór fram á skólatíma í einni kennslustund og voru tveir rannsakendur viðstaddir ásamt 15 nemendum. Engum persónuupplýsingum var safnað við fram-kvæmd viðtalsins.

Undirbúningsmæling II. Leyfi var fengið hjá skólastjórnendum í skóla í Reykjavík til að fá nemendur til að svara endurbættri útgáfu spurningalistans. Gagnasöfnunin fór fram samtímis í tveimur kennslustofum í einni kennslustund. Tveir rannsakendur skiptu með sér verkum og lásu upp staðlaðan fyrirlagnartexta við upphaf kennslu-stundarinnar. Bekkjunum var stillt upp eins og um próf væri að ræða. Engum persónu-upplýsingum var safnað við fyrirlögnina.

Forprófun. Leyfi var fengið hjá skóla-stjórnendum í skóla í Reykjavík til að fá

nemendur til að svara lokaútgáfu spurninga-listans í heild. Gagnasöfnunin fór fram samtímis í þremur kennslustofum í einni kennslustund. Þrír rannsakendur skiptu með sér verkum og lásu upp staðlaðan fyrirlagnartexta við upphaf kennslustundarinnar. Bekkjunum var stillt upp eins og um próf væri að ræða. Engum persónuupplýsingum var safnað við fyrirlögnina.

Langtímarannsókn. Rannsóknin var tilkynnt til Persónuverndar undir númerinu S5799. Skriflegu samþykki var safnað frá skólum og foreldrum allra barna sem tóku þátt í rannsókninni. Gagnasöfnunin fór fram í október og apríl skólaárið 2012-2013. Tímasetning fyrirlagnanna í hverjum skóla var ákveðin í samráði við kennara og skólastjórnendur á hverjum stað. Tveir þjálfaðir rannsakendur fóru í viðkomandi skóla og framkvæmdu fyrirlögnina í samráði við umsjónarkennara á hverjum stað. Hverjum bekk var stillt upp eins og um próf væri að ræða og lesinn var upp staðlaður texti um tilgang og framkvæmd rannsóknarinnar. Rannsakendurnir svöruðu einnig spurningum nemenda með samræmdum hætti eftir því sem aðstæður leyfðu. Reynt var að fá nemendur sem voru veikir, fluttir eða í löngu leyfi til að svara spurningalistanum með hjálp umsjónarkennara og/eða foreldra.

Mælitæki

SOC (tvíkosta framsetning). SOC mælingin samanstendur af 18 tvíkosta atriðum þar sem svarandi þarf að velja hvort hann líkist frekar manneskju A eða B. Listinn var upprunalega þýddur og bakþýddur úr ensku yfir á íslensku af rannsakendum með ítarlega þekkingu á hugtakaskilgreiningu (e. *conceptual definition*) sjálfstjórnunar (Steinunn Gestsdóttir o.fl., 2011). Hvert atriði lýsir annarsvegar hegðun sem er dæmigerð fyrir sjálfstjórnun (sjá viðauka) og hinsvegar hegðun sem telst andstæða sjálfstjórnunar. SOC mælingin mælir meðvitaða sjálfstjórnun sem snýr að þeim markmiðum sem fólk setur sér í lífinu. Mælingin, eins og lýst er í inngangi, er fjórþætt

og inniheldur spurningar um hversu vel fólki gengur að velja sér markmið (val), hversu vel því gengur að finna leiðir að markmiðum sínum (hámörkun), hversu vel því gengur að skipta út leiðum ef þörf krefur (uppbót) og hversu vel því gengur að skipta um markmið ef fyrri markmið reynast óraunhæf (val byggt á tapi). Fjórða þættinum (val byggt á tapi) var sleppt þar sem sá þáttur er ekki talinn eiga við unglingsárin sem einkennast af vexti á flestum sviðum þroska. Dæmi um atriði sem lýsir því hversu vel fólki gengur að finna leiðir að markmiðum sínum er: „Ég íhuga vandlega hvernig ég get best náð markmiðum mínum."

SOC (fyrri Likert útgáfa). Í samræmi við niðurstöður úr þáttagreiningu á tvíkosta og Likert framsetningu SOC mælitækisins (Geldhof, Bowers, Gestsdottir, Napolitano og Lerner, í prentun; Geldhof o.fl., 2012) var útbúin SOC mæling með fimm Likert svarkostum (Mjög miklu leyti, Miklu leyti, Svolitlu leyti, Litlu leyti, Alls engu leyti). Mælingin samanstóð af 18 atriðum þar sem hvert atriði lýsti einungis hegðun sem er dæmigerð fyrir sjálfstjórnun (sjá viðauka). Um var að ræða sömu atriðin og notuð voru til að lýsa dæmigerðri sjálfstjórnun í tvíkosta framsetningunni sem lýst er hér á undan. Þetta mælitæki var einungis notað í undirbúningsmælingu I.

SOC (seinni Likert útgáfa). Að lokinni undirbúningsmælingu I og rýnihópsviðtali með 14 ára unglingum var útbúin önnur Likert útgáfa af SOC mælitækinu (sjá viðauka) með nýjum svarkostum (Mjög líkt mér, Frekar líkt mér, Hvorki líkt né ólíkt mér, Frekar ólíkt mér og Mjög ólíkt mér). Mælitækið samanstóð af 18 atriðum þar sem hvert atriði lýsti hegðun sem er dæmigerð fyrir sjálfstjórnun. Um var að ræða sömu atriðin og notuð voru í fyrri Likert útgáfunni sem lýst er hér á undan.

Sjálfstjórnun í námi. Mæling á sjálfstjórnun í námi var á öllum stigum notuð til að meta samleitniréttmæti við SOC mælingarnar.

Mælingin hefur verið aðlöguð til notkunar með unglingum á Íslandi samkvæmt leiðbeiningum frá Albert Bandura (2006). Mælingin samanstendur af sjö atriðum. Nemendur eru beðnir um að gefa til kynna á fimm punkta Likert kvarða hversu vel þeir treysta sér til að geta til dæmis „haldið sér að námi þegar eitthvað annað áhugavert er í boði". Mælitækið samanstendur af sjö atriðum með svarmöguleikunum: Get ekki, Get eiginlega ekki, Get nokkurn veginn, Get eiginlega alveg, Get alveg. Fjöldi mælinga hefur staðfest áreiðanleika mælitækisins þar sem Cronbachs alfa hefur mælst á bilinu 0,78 til 0,84 (Usher og Pajares, 2008).

Tölfræðileg úrvinnsla

Lýsandi tölfræði, mat á áreiðanleika og samleitniréttmæti fóru fram með tölfræðiforritinu SPSS 20. Áreiðanleiki SOC mælitækisins var metinn með Cronbachs alfa áreiðanleikastuðlinum. Miðað var við að tölugildi á bilinu 0,7-0,8 gæfi til kynna ásættanlegan áreiðanleika (Kline, 1999). Samleitniréttmæti SOC var metið með Pearsons fylgnistuðlinum. Miðað var að jákvæð fylgni allt að 0,85 gæfi til kynna samleitniréttmæti en fylgni hærri en 0,85 gæfi til kynna að mælitækin væru að mæla sömu hugsmíðina (Kline, 2011).

Mat á hugtakaréttmæti SOC mælitækisins (seinni Likert útgáfa) fór fram með staðfestandi þáttagreiningu í tölfræðiforritinu Mplus 7.1. Stærsti hluti þeirra tölfræðiaðferða sem í boði eru í dag byggja á kenningunni um miðsækni og gera ráð fyrir normaldreifingu svara. Þar sem lýsandi tölfræði úr forprófun mælitækisins gaf til kynna að niðurstöður SOC mælitækisins væru normaldreifðar fór mat á stuðlum formgerðarlíkans SOC fram með sennileikamatsaðferðinni (e. *maximum likelihood estimation method*). Til að eiga möguleika á samanburði við fyrri rannsóknir á unglingum voru metin líkön sem gerðu ráð fyrir þremur undirliggjandi þáttum (vali, hámörkun og uppbót) og einnig einum almennum sjálfstjórnunarþætti. Þáttahleðsla fyrsta atriðis

í hverjum undirþætti var skilgreind sem tölugildið einn til að unnt væri að meta líkanið. Fylgni á milli þátta var ekki takmörkuð og mælivilla hverrar spurningar var einnig metin án takmarkana. Ekki var gert ráð fyrir að fylgni væri á milli dreifni villuliða milli ólíkra atriða.

Mátgæði (e. *goodness of fit*) þriggja þátta líkans voru metin út frá fleiri en einni tegund mátstuðla. Lagt var mat á SRMR (e. *standardized root mean squared residual*) stuðul, kí-kvaðrat stuðul, CFI (e. *comparative fit index*) stuðul og RMSEA (e. *root means square error*) stuðul. SRMR stuðullinn tekur gildi frá núll til einn og því lægri sem stuðullinn er því betur fellur líkanið að gögnunum. Litið var svo á að líkanið félli vel að gögnunum ef stuðullinn var 0,08 eða lægri. RMSEA stuðullinn getur tekið gildi á bilinu núll til einn en því lægri sem hann er því betur fellur líkanið að gögnunum. Litið var svo á að líkanið félli vel að gögnunum ef RMSEA stuðullinn var 0,06 eða lægri. CFI stuðullinn tekur gildi á bilinu núll til einn en því hærri sem hann er því betur fellur líkanið að gögnunum. Litið var svo á að líkanið félli vel að gögnunum ef CFI stuðullinn var 0,95 eða hærri (West, Taylor og Wu, 2013).

Eftir að mat hafði verið lagt á mátgæði líkansins var tekið tillit til fylgni á milli þátta, þáttahleðslna og breytingastuðla til að ákveða hvort breytingar á líkaninu væru æskilegar.

Niðurstöður

Söfnun gagnanna fór fram í nokkrum skrefum svo unnt væri að þróa SOC mælitækið og um leið svara þeim rannsóknarspurningum sem settar voru fram. Grein er gerð fyrir niðurstöðum rannsóknarinnar í sömu röð og söfnun gagna fór fram. Eins og fram kemur síðar í þessum kafla kom fjölþátta formgerð SOC mælitækisins ekki fram meðal unglinga. Af þeim sökum er ekki fjallað um S, O, og C þættina, heldur eru niðurstöðurnar takmarkaðar við einsþátta formgerð mælitækisins.

Undirbúningsmæling I

Niðurstöðurnar úr undirbúningsmælingu I voru notaðar til að bera saman tvíkosta- og Likert framsetningu SOC mælitækisins. Áreiðanleikagreining sýndi að áreiðanleiki níu atriða SOC mælingar fór frá því að vera næstum viðunandi ($\alpha = 0,67$) með tvíkosta framsetningunni yfir í að vera viðunandi með því að nota Likert framsetninguna ($\alpha = 0,77$). Niðurstöður áreiðanleikagreiningarinnar voru staðfestar í opnum svörum þar sem nemendur voru hvattir til að koma með athugasemdir varðandi framsetningu mælitækja í könnuninni. Í heildina gerði 21 (15%) nemandi athugasemd. Af þessum hópi gáfu 15 nemendur (71%) til kynna að erfitt væri að skilja og nota tvíkosta framsetningu SOC mælingarinnar í samanburði við Likert framsetninguna. Þeir sex nemendur sem ekki minntust á SOC mælitækin komu með athugasemdir við aðrar mælingar eða komu með athugasemdir sem ekki tengdust efni spurningalistans.

Báðar útgáfur mælitækisins sýndu fylgni við sjálfstjórnun í námi. Styrkur sambandsins fór frá því að vera jákvæður og sterkur ($r(128) = 0,43$, $p < 0,001$) yfir í að vera jákvæður og mjög sterkur með því að nota Likert útgáfuna ($r(138) = 0,49$, $p < 0,001$). Skekkjan í dreifingu svaranna í níu atriða útgáfunni reyndist ásættanleg (-0,28) með notkun Likert framsetningarinnar en ris dreifingarinnar reyndist hinsvegar óásættanlegt (1,31).

Rýnihópsviðtal

Í ljósi þess að Likert framsetning á SOC mælitækinu gaf í senn áreiðanlegri og réttmætari mælingu á sjálfstjórnun og einnig með tilliti til þess að háskólanemar áttu í erfiðleikum með tvíkosta framsetningu þess var ákveðið að vinna áfram með Likert framsetninguna til notkunar með unglingum. Rýnihópur 15 unglinga í 9. bekk í skóla á Reykjavíkursvæðinu var fenginn til að svara og ræða Likert framsetningu mælitækisins með það fyrir augum að gera mælitækið skiljanlegra. Á meðan á umræðunni stóð bentu nokkrir unglinganna á að erfitt væri að skilja

1. tafla. Sjálfstjórnunaratriði SOC sem eftir standa eftir staðfestandi þáttagreiningu meðal 14 og 18 ára unglinga á Íslandi og meðal 11 til 13 ára bandarískra unglinga.

	Fullyrðing	Ísland 14 ára	14* ára	18* ára	Bandaríkin 11# ára	12# ára	13+ ára
	Atriði kvarðans Val (e. *Selection*)						
m)	Þegar ég set mér markmið þá stend ég við það	x	x	x	x	x	x
r)	Ég vinn alltaf að einu markmiði í einu	x	x	x	x	x	x
a)	Ég einbeiti mér að fáum hlutum í einu						
b)	Ég velti því mikið fyrir mér hvað er mér mikilvægt						
k)	Ég stefni alltaf að einu, mikilvægu markmiði í einu						
l)	Þegar ég hef virkilega ákveðið hvað ég vil í lífinu þá vinn ég að einu eða tveimur mikilvægum markmiðum						
	Atriði kvarðans Hámörkun (e. *Optimization*)						
c)	Til þess að ná markmiði mínu reyni ég eins margar nýjar leiðir eins og á þarf að halda	x	x	x	x	x	x
f)	Þegar ég ætla mér að ná erfiðu markmiði, bíð ég eftir réttu augnabliki og besta tækifærinu						
g)	Ég íhuga vandlega hvernig ég get best náð markmiðum mínum	x		x	x	x	x
h)	Ég geri allt hvað ég get til að ná markmiðum mínum	x	x	x	x	x	x
j)	Þegar ég byrja á einhverju sem mér finnst skipta máli en veit að verður erfitt, þá legg ég sérstaklega hart að mér	x	x	x	x	x	x
n)	Þegar ég vil ná árangri, þá skoða ég líka hvernig aðrir hafa gert það						
	Atriði kvarðans Uppbót (e. *Compensation*)						
e)	Þegar mér finnst eitthvað vera mikilvægt, þá velti ég fyrir mér hvort ég þurfi að verja meiri tíma í það eða að vera duglegri	x		x	x	x	x
o)	Þegar hlutirnir ganga ekki jafn vel og áður reyni ég að gera þá öðruvísi	x	x	x	x	x	x
q)	Þegar eitthvað gengur ekki jafn vel og venjulega athuga ég hvernig aðrir hafa farið að	x	x	x	x	x	x
d)	Þegar eitthvað virkar ekki eins vel og áður, fæ ég ráðleggingar á netinu eða les bók						
i)	Þegar hlutirnir ganga illa, þigg ég hjálp frá öðrum						
p)	Þegar ég get ekki leyst verkefni eins vel og áður, bið ég einhvern annan að gera það fyrir mig						

*Niðurstöður úr Steinunn, Sigrún og Fanney (2011)
#Niðurstöður úr Gestsdottir og Lerner (2007)
+Niðurstöður úr Zimmerman o.fl. (2007)

svarmöguleikana: Mjög miklu leyti, Miklu leyti, Svolitlu leyti, Litlu leyti og Alls engu leyti. Nemendurnir mæltu með notkun á algengari tegund svarmöguleika: Mjög líkt mér, Frekar líkt mér, Hvorki líkt né ólíkt mér, Frekar ólíkt mér og Mjög ólíkt mér. Svarmöguleikarnir sem rýnihópurinn mælti með reyndust meira viðeigandi en fyrri svarmöguleikar þegar þeir voru bornir saman við skýringartextann sem fylgdi upprunalegu tvíkosta SOC mælingunni (Baltes, Baltes, Freund og Lang, 1999).

Undirbúningsmæling II

Önnur undirbúningsmæling var framkvæmd til að athuga áhrif þess að breyta svarmöguleikunum frá „Mjög miklu leyti, Miklu leyti, Svolitlu… o.s.frv." yfir í „Mjög líkt mér, Frekar líkt mér, Hvorki … o.s.frv". Niðurstöðurnar bentu til þess að breytingin á svarmöguleikunum leiddi til normaldreifingar svara þar sem óásættanlegt ris dreifingarinnar úr undirbúningsmælingu I fór frá því að vera 1,31 niður í að vera 0,59 á sama tíma og skekkja dreifingarinnar hélst ásættanleg (0,45). Normaldreifing svara úr sálfræðilegum mælitækjum er forsenda fyrir stærstum hluta þekktra aðferða í tölfræði og því var ákveðið að nota seinni útgáfu svarmöguleikanna í forprófun spurningalistans í heild sinni.

Forpróf

Að loknu rýnihópsviðtali og tveimur undirbúningsmælingum var lokaútgáfa spurningalistans forprófuð. Lokaútgáfa spurningalistans innihélt, auk fjölda annarra mælitækja, SOC mælitæki með 18 atriðum sem lýsa dæmigerði hegðun fyrir sjálfstjórnun með fimm svarmöguleikum: Mjög líkt mér, Frekar líkt mér, Hvorki líkt né ólíkt mér, Frekar ólíkt mér og Mjög ólíkt mér. Niðurstöðurnar sýndu ásættanlegan áreiðanleika fyrir níu atriða útgáfu mælitækisins ($\alpha = 0,77$).

Langtímarannsókn

Gögn úr fyrstu mælingu af fjórum í langtímarannsókninni sem þessi rannsókn er hluti af (N = 539) gáfu nægilegt afl til að meta

formgerð SOC mælingarinnar og þar með hugtakaréttmæti mælitækisins. Í ljós kom að fyrsta líkanið, þar sem metið var þriggja þátta líkan, féll illa að gögnunum ($\chi^2 = 574,72$, $p = 0,00$; SRMR = 0,066; RMSEA = 0,082 (CI = 0,075-0,089); CFI = 0,82). Fylgni á milli þáttanna þriggja var mjög há ($r = 0,79$-0,98), sem gaf einnig til kynna að rangt líkan væri fellt að gögnunum. Sérstaka athygli vakti að fylgnin milli þess að finna leiðir að markmiðum (hámörkun) og að skipta út markmiðum (uppbót) var sérstaklega há ($r = 0,98$). Í ljósi þess hve fylgnin var há milli þáttanna þriggja var ákveðið að líkan númer tvö gerði ráð fyrir einum níu atriða sjálfstjórnunarþætti í stað þriggja þátta. Er það í samræmi við það sem erlendar rannsóknir með bandarískum unglingum hafa mælt með (Geldhof o.fl., í prentun; Gestsdottir o.fl., 2009; Zimmerman o.fl., 2007).

Þegar einungis hin níu hefðbundnu atriði voru mátuð við gögnin kom í ljós að einsþátta líkan með níu atriðum féll ásættanlega að gögnunum ($\chi^2 = 70,03$, $p = 0,00$; SRMR = 0,035; RMSEA = 0,06 (CI = 0,040-0,072); CFI = 0,95). Áreiðanleiki kvarðans í fyrstu gagnasöfnun reyndist ásættanlegur ($\alpha = 0,75$). Að lokum var einsþátta, níu atriða mælingin mátuð við gögn úr annarri gagnasöfnun langtímarannsóknarinnar sem endurtekin var sex mánuðum eftir þá fyrri. Í ljós kom að mátgæðin fóru frá því að vera góð yfir í að vera mjög góð við endurtekninguna ($\chi^2 = 58,34$, $p = 0,0004$; SRMR = 0,032; RMSEA = 0,05 (CI = 0,031-0,065); TLI = 0,96; CFI =0,97) (West, Taylor og Wu, 2013). Áreiðanleiki kvarðans við aðra gagnasöfnun reyndist einnig góður ($\alpha = 0,84$). Í töflu 1 er að finna þau atriði sem mynduðu SOC sjálfstjórnunarmælitækið í þessari rannsókn og nokkrum skyldum rannsóknum merkt með X.

Umræður

Markmið þessarar rannsóknar var tvíþætt. Í fyrsta lagi var athugað hvort hægt væri að bæta sjálfstjórnunarmælitækið SOC með því

að setja svarmöguleika mælitækisins fram á Likert kvarða. Í öðru lagi var kannað hvort breytt framsetning mælitækisins yrði til þess að hægt væri að staðfesta eins- eða fjölþátta formgerð mælitækisins.

Þrátt fyrir ýmsa kosti sem komu fram við Likert framsetningu á SOC mælitækinu sýndi staðfestandi þáttagreining á gögnum úr fyrstu gagnasöfnun ekki þriggja þátta formgerð. Það er ólíkt því sem sýnt hefur verið fram á meðal fullorðinna. Níu atriða einsþátta formgerð mælitækisins var staðfest meðal 14 ára ungmenna á Íslandi að hausti í 9. bekk og aftur sex mánuðum síðar sem bendir til hugtakaréttmætis mælitækisins við íslenskar aðstæður. Fyrri rannsóknir á tvíkosta framsetningu SOC meðal 14 ára ungmenna á Íslandi hafa ekki staðfest níu atriða formgerð mælitækisins (sjá Steinunn Gestsdóttir o.fl., 2011) og má vera að Likert framsetning af því tagi sem mælt er með í þessari rannsókn eigi betur við en tvíkosta framsetning þegar mælitækið er notað með íslenskum ungmennum. Endurtekin staðfesting á formgerð níu atriða útgáfunnar sex mánuðum eftir fyrstu mælingu bendir enn frekar til þess að mælitækið búi yfir langtímastöðugleika. Í heild má segja að niðurstöður rannsóknarinnar bendi til þess að framsetning SOC atriðanna á Likert kvarða hafi jákvæð áhrif á áreiðanleika og réttmæti mælitækisins í samanburði við fyrri rannsóknir á tvíkosta framsetningu.

Niðurstöðurnar gefa til kynna að íslenskir unglingar geri lítinn greinarmun á ólíkum undirþáttum SOC mælingarinnar ólíkt því sem staðfest hefur verið í rannsóknum meðal fullorðinna. Sérstaka athygli vakti að unglingarnir gerðu nær engan greinarmun á því að finna áhrifaríkar leiðir að markmiðum (hámörkun) og að skipta út markmiðum (uppbót). Þessar niðurstöður eru mikilvægar fyrir þá sem vinna með styrkingu sjálfstjórnunar meðal unglinga þar sem þær gefa til kynna að unglingar eigi erfiðara með að taka afstöðu til ólíkra leiða að þeim markmiðum sem þeir setja sér en fullorðnir. Ástæður þessa má mögulega rekja til ólíkrar reynslu, vitsmunaþroska

eða áhugahvatar unglinga í samanburði við fullorðna. Niðurstöðurnar benda til þess að gagnlegt geti verið að vinna skipulega með ólíkar leiðir að algengum markmiðum, til dæmis í skólastarfi, til að víkka reynsluheim og sjálfsþekkingu unglinga. Frekari rannsókna er þörf til að kanna hvaða aðferðir reynast best við styrkingu sjálfstjórnunar á unglingastigi.

Fjöldi nýlegra erlendra rannsókna er fyrirliggjandi þar sem einsþátta, níu atriða SOC mæling hefur verið notuð til að spá fyrir um ýmsa þætti í farsælum þroska ungmenna. Staðfesting á einsþátta, níu atriða útgáfu meðal unglinga á Íslandi gefur því tækifæri á samanburði við fyrri rannsóknir sem færir út mörk þekkingar á sjálfstjórnun ungmenna á Íslandi.

Mælingar á sjálfstjórnun með notkun spurningalista eins og hér er gert ber að taka með fyrirvara þar sem sjálfsmat gefur ekki fullnægjandi mynd af raunverulegri getu til sjálfstjórnunar. Hugmyndir nemenda um að hvaða markmiðum er samfélagslega viðurkennt að keppa og hversu vel nemandinn samsamar sig þeim markmiðum gætu haft áhrif á hugmyndir hans um eigin sjálfstjórnun. Gagnleg viðbót við núverandi mælingar væri ef síðari rannsóknir myndu þróa og nota í meira mæli beinar mælingar á hegðun við mat á meðvitaðri sjálfstjórnun, til dæmis með rannsóknum þar sem hegðun unglinga á vettvangi er skipulega skráð af þjálfuðum rannsakendum. Mælingar úr taugasálfræði á þroska og virkni svæða í heilanum sem talin eru spila lykilhlutverk í sjálfstjórnun væru einnig gagnleg leið til að margprófa og staðfesta réttmæti þeirra mælitækja sem nú eru til staðar.

SOC mælitækið á, eins og áður segir, uppruna sinn í öldrunarfræðum í Þýskalandi. Þær rannsóknir sem liggja að baki mælitækinu byggja á svörum þátttakenda sem spanna mjög breitt aldursbil eða 71 ár (Freund og Baltes, 2002). Þær rannsóknir sem vísað hefur verið til í þessari grein innihalda einungis svör ungs fólks. Mögulegt er að atriði mælitækisins greini ekki á milli undirþátta SOC á fullnægjandi hátt

þegar einungis er um unga þátttakendur að ræða. Ekki er ljóst af hverju ungmenni greina ekki á milli mismunandi sjálfstjórnunarferla líkt og eldra fólk. Þetta þyrfti að skoða nánar og mögulega endurskoða SOC atriðin með það takmark í huga að aðgreina með skýrari hætti þær hugsmíðar sem SOC kenningin gerir ráð fyrir í svörum ungra þátttakenda. Í því sambandi væri æskilegt að beita eigindlegum aðferðum til að fá innsýn í þær leiðir sem ungmennum þykja árangursríkastar í að setja sér markmið og ná þeim.

Niðurstöður úr fyrri undirbúnings-mælingunni meðal háskólanema bentu til þess að fullorðnu fólki finnist erfitt að svara SOC atriðunum þegar þau eru sett fram með tvíkosta framsetningu. Nýleg sænsk rannsókn hefur jafnframt gefið til kynna að atriði a og k (sjá 1. töflu) hafa neikvæð áhrif á áreiðanleika undirþáttarins val meðal fullorðinna í Svíþjóð (Viglund o.fl., 2013). Fjölþáttaformgerð SOC mælitækisins hefur jafnframt ekki komið fram í tveimur rannsóknum á 18 ára nemendum í Bandaríkjunum og á Íslandi (Geldhof o.fl., 2012; Steinunn Gestsdóttir o.fl., 2011). Niðurstöður þessarar og fyrri rannsókna benda því til þess að nauðsynlegt sé að skoða formgerð SOC mælitækisins nánar meðal fullorðinna.

Með þeim breytingum sem lagðar eru til á framsetningu SOC mælitækisins í þessari grein eru stigin skref í átt að því að hanna mælitæki á meðvitaðri sjálfstjórnun sem hæfir ungmennum. Nú liggur fyrir einsþátta útgáfa SOC mælitækisins með Likert svarmöguleikum sem er áreiðanleg og réttmæt í notkun með unglingum við upphaf og lok 9. bekkjar á Íslandi og gefur sú staðreynd því möguleika á frekari rannsóknum á meðvitaðri sjálfstjórnun meðal unglinga á Íslandi.

Kristján Ketill Stefánsson is a Ph.D. student at the Faculty of Education Studies, University of Iceland. Steinunn Gestsdóttir is an Associate Professor at the Faculty of Psychology, University of Iceland. Sigur-grímur Skúlason is a Department Manager at the Ed-ucation Testing Institute. This study was supported by the University of Iceland Doctoral Grant and the University of Iceland Scholarship Fund. Correspon-dence concerning this article should be addressed to Kristján Ketill Stefánsson. E-mail: kristjan@hi.is

The development and validation of the SOC scale for youth in Iceland

Intentional self-regulation refers to people's ability to set goals and find effective means to achieve their goals. The SOC (selection, optimization, compensation) measure has been widely used in measuring intentional self-regulation among adults. Research has indicated the importance of SOC abilities during adolescence but a reliable and valid SOC measure for use with adolescents has not been available. The goal of this research was twofold. First, the means of improving the reliability of the SOC measure were explored by changing the scale of the measure from a forced choice scale to a Likert scale. Second, a confirmatory factor analysis (CFA) was performed to investigate the construct validity of SOC when using Likert answer options. A pilot study with college students, and adolescents from a secondary school in Reykjavík, indicated that a Likert scale representation of the SOC items produced higher reliability and more convergent validity then a forced choice representation. A CFA on data from 539 students in 9th grade (46% girls, mean age 14.3 years) showed that a tripartite structure of SOC did not fit the data adequately. However, a single factor structure using nine items from the SOC questionnaire indicated a good fit as has been shown by previous research with adolescents in other cultures. These results indicate that a single factor nine item SOC measure with Likert answer options gives a reliable and valid measure of intentional self-regulation at the beginning and at the end of ninth grade in Iceland.

Keywords: Self-regulation, SOC, confirmatory factor analysis, measurement, adolescents.

Heimildir

Baltes, P. (1997). On the incomplete architecture of human ontogeny - Selection, optimization, and compensation as foundation of developmental theory. *American Psychologist, 52*(4), 366-380. doi: 10.1037//0003-066X.52.4.366

Baltes, P., Baltes, M., Freund, A. og Lang, F. (1999). *The measurement of selection, optimization, and compensation (SOC) by self report. Technical report.* Berlín, Þýskaland: Max-Planck-Institut für Bildungsforschung.

Bandura, A. (2006). Guide for constructing self-efficacy scales. Í F. Pajares og T. Urdan (Ritstj.), *Adolescence and education* (bls. 306-337). Greenwich, CT: IAP - Information Age.

Bowers, E. P., Gestsdottir, S., Geldhof, G. J., Nikitin, J., von Eye, A. og Lerner, R. M. (2011). Developmental trajectories of intentional self regulation in adolescence: The role of parenting and implications for positive and problematic outcomes among diverse youth. *Journal of Adolescence, 34*(6), 1193-1206. doi: 10.1016/j.adolescence.2011.07.006

Brandstädter, J. (2006). Action perspectives on human development. Í W. Damon og R. M. Lerner (Ritstj.), *Handbook of child psychology: Theoretical models of human development* (6. útgáfa, bls. 516-568). New York, NY: Wiley. af http://www.loc.gov/catdir/enhancements/fy0643/2005043951-d.html

Brody, G. H. og Ge, X. (2001). Linking parenting processes and psychological functioning and alcohol use during early adolescence. *Journal of Family Psychology, 15*, 82–94.

Freund, A. M. og Baltes, P. B. (2002). Life-management strategies of selection, optimization, and compensation: Measurement by self-report and construct validity. *Journal of Personality and Social Psychology, 82*(4), 642-662. doi: 10.1037//0022-3514.82.4.642

Geldhof, G. J., Bowers, E. P., Gestsdottir, S., Napolitano, C. M. og Lerner, R. M. (í prentun). Self-regulation across adolescence: exploring the structure of selection, optimization, and compensation. *Journal of Research in Adolescence.*

Geldhof, G. J., Little, T. D. og Hawley, P. H. (2012). Two measures of self-regulation for young adults and late adolescents in the academic and social domains. *International Journal of Behavioral Development, 36*(6), 476-488. doi: 10.1177/0165025412462153

Gestsdottir, S., Geldhof, G. J., Paus, T., Freund, A. M., Adalbjarnardottir, S., … Lerner, R. M. (2014). Self-regulation processes among youth in four western cultures: Is there an adolescent-specific structure of the Selection-Optimization-Compensation (SOC) model? *Handrit sent til birtingar.*

Gestsdottir, S. og Lerner, R. M. (2007). Intentional self-regulation and positive youth development in early adolescence: Findings from the 4-H study of positive youth development. *Developmental Psychology, 43*(2), 508-521. doi: 10.1037/0012-1649.43.2.508

Gestsdottir, S. og Lerner, R. M. (2008). Positive development in adolescence: The development and role of intentional self-regulation. *Human Development, 51*(3), 202-224. doi: 10.1159/000135757

Gestsdottir, S., Lewin-Bizan, S., von Eye, A., Lerner, J. V. og Lerner, R. M. (2009). The structure and function of selection, optimization, and compensation in middle adolescence: Theoretical and applied implications. *Journal of Applied Developmental Psychology, 30*(5), 585-600. doi: 10.1016/j.appdev.2009.07.001

Kline, P. (1999). *The handbook of psychological testing.* London: Routledge.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3. útgáfa). New York: Guilford Press.

Lerner, J. V., Phelps, E., Forman, Y. og Bowers, E. P. (2009). Positive Youth Development. Í R. M. Lerner og L. D. Steinberg (Ritstj.), *Handbook of adolescent psychology* (3. útgáfa). Hoboken, NJ: Wiley.

Massey, E. K., Gebhardt, W. A. og Garnefski, N. (2008). Adolescent goal content and pursuit: A review of the literature from the past 16 years. *Developmental Review, 28*(4), 421-460. doi: 10.1016/J.Dr.2008.03.002

McClelland, M. M., Acock, A. C. og Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly, 21*, 471-490.

McClelland, M. M. og Cameron, C. E. (2011). *Self-Regulation and Academic Achievement in Elementary School Children* (Nr. 1534-8687). doi: 10.1002/Cd.302

McClelland, M. M., Morrison, F. J. og Holmes, D. L. (2000). Children at-risk for early academic problems: The role of learning-related social skills. *Early Childhood Research Quarterly, 15*, 307-329. doi: 10.1016/S0885-2006(00)00069-7

McClelland, M. M., Ponitz, C. C., Messersmith, E. E. og Tominey, S. (2010). Self-regulation: Integration of cognition and emotion. Í R. M. Lerner (Ritstj.), *The Handbook of Life-Span Development* (bls. 509-554). New Jersey, NJ: John Wiley & Sons. doi: 10.1002/9780470880166.hlsd001015

Mennta- og menningarmálaráðuneytið. (2011). *Aðalnámskrá grunnskóla - almennur hluti.* Sölvhólsgötu 4: Mennta- og menningarmálaráðuneytið.

Overton, W. F. (2013). *A New Paradigm for Developmental Science: Relationism and Relational-Developmental Systems. Applied Developmental Science, 17*(2), 94-107. doi: 10.1080/10888691.2013.778717

Quinn, P. D. og Fromme, K. (2010). Self-regulation as a protective factor against risky drinking and sexual behavior. *Psychology of Addictive Behaviors, 24,* 376–385. doi: 10.1037/a0018547

Steinunn Gestsdóttir og Lerner, R. M. (2007). Hlutverk sjálfstjórnar í æskilegum þroska barna og unglinga. *Sálfræðiritið, 12,* 35-53.

Steinunn Gestsdóttir, Sigrún Aðalbjarnardóttir og Fanney Þórsdóttir. (2011). *Formgerð sjálf-stjórnar: rannsókn meðal íslenskra ungmenna. Sálfræðiritið, 16,* 7-21.

Usher, E. L. og Pajares, F. (2008). Self-efficacy for self-regulated learning - A validation study. *Educational and Psychological Measurement, 68*(3), 443-463. doi: 10.1177/0013164407308475

Viglund, K., Alex, L., Jonsen, E., Lundman, B., Norberg, A., … Nygren, B. (2013). Psychometric properties of the Swedish version of the selection, optimization, compensation questionnaire. *Scandinavian Journal of Caring Sciences, 27*(2), 460-467. doi: 10.1111/j.1471-6712.2012.01027.x

West, S., Taylor, A. og Wu, W. (2013). Model fit and model selection in structural equation modeling. Í R. H. Hoyle (Ritstj.), *Handbook of structural equation modeling.* NY: Guilford.

Wiese, B. S., Freund, A. M. og Baltes, P. B. (2000). Selection, optimization, and compensation: An action-related approach to work and partnership. *Journal of Vocational Behavior, 57*(3), 273-300. doi: 10.1006/jvbe.2000.1752

Wigfield, A. og Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68-81. doi: 10.1006/ceps.1999.1015

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*(2), 64-70.

Zimmerman, S. M., Phelps, E. og Lerner, R. M. (2007). Intentional self-regulation in early adolescence: Asessing the structure of selection, optimization, and compensation processes. *International Journal of Developmental Science, 1*(3), 272-299. doi: 10.3233/dev-2007-1310

Viðauki

Átján atriða SOC spurningalistinn, einungis svarmöguleikar sem lýsa sjálfstjórnun (Freund og Baltes, 2002).

Val
1. Ég einbeiti mér að fáum hlutum í einu
2. Ég stefni alltaf að einu, mikilvægu markmiði í einu
3. Ég vinn alltaf að einu markmiði í einu
4. Þegar ég hef virkilega ákveðið hvað ég vil í lífinu þá vinn ég að einu eða tveimur mikilvægum markmiðum
5. Þegar ég set mér markmið þá stend ég við það
6. Ég velti því mikið fyrir mér hvað er mér mikilvægt

Hámörkun
1. Ég geri allt hvað ég get til að ná markmiðum mínum
2. Til þess að ná markmiði mínu reyni ég eins margar nýjar leiðir eins og á þarf að halda
3. Þegar ég ætla mér að ná erfiðu markmiði bíð ég eftir réttu augnabliki og besta tækifærinu
4. Þegar ég byrja á einhverju sem mér finnst skipta máli en veit að verður erfitt, þá legg ég sérstaklega hart að mér
5. Þegar ég vil ná árangri, þá skoða ég líka hvernig aðrir hafa gert það
6. Ég íhuga vandlega hvernig ég get best náð markmiðum mínum

Uppbót
1. Þegar eitthvað gengur ekki jafn vel og venjulega athuga ég hvernig aðrir hafa farið að
2. Þegar eitthvað virkar ekki eins vel og áður, fæ ég ráðleggingar á netinu eða les bók
3. Þegar mér finnst eitthvað vera mikilvægt þá velti ég fyrir mér hvort ég þurfi að verja meiri tíma í það eða að vera duglegri
4. Þegar hlutirnir ganga illa, þigg ég hjálp frá öðrum
5. Þegar hlutirnir ganga ekki jafn vel og áður reyni ég að gera þá öðruvísi
6. Þegar ég get ekki leyst eitthvað verkefni eins vel og áður þá bíð ég einhvern annan að gera það fyrir mig

Svarmöguleikar: Mjög líkt mér, Frekar líkt mér, Hvorki líkt né ólíkt mér, Frekar ólíkt mér, Mjög ólíkt mér.

Paper II: Stefansson, Kristjan. K., Gestsdottir, S., Geldhof, G. J., Skulason, S., & Lerner, R. M. (2015). A bifactor model of school engagement: Assessing general and specific aspects of behavioral, emotional and cognitive engagement among adolescents. *International Journal of Behavioral Development*, **1–10. doi:10.1177/0165025415604056**

A Bifactor Model of School Engagement: Assessing General and Specific Aspects of Behavioral, Emotional and Cognitive Engagement among Adolescents

Kristjan K. Stefansson,[1] **Steinunn Gestsdottir,**[1] **G. John Geldhof,**[2] **Sigurgrimur Skulason,**[3] **and Richard M. Lerner**[4]

Abstract

School engagement involves cognitive, emotional, and behavioral components that overlap conceptually. This conceptual ambiguity has led to measures that have either consisted of one general factor or separate correlated factors. However, neither approach can sufficiently account for both the uniqueness and the overlap of the subcomponents. The bifactor model has been recommended to determine the degree to which a measure is unidimensional versus multidimensional. In this study, we examined the validity of a multidimensional measure of school engagement in adolescence, the Behavioral-Emotional-Cognitive School Engagement Scale (BEC-SES; Li & Lerner, 2013), by comparing the model fit and predictive power of the widely-used one- and three-factor models with a bifactor model. Using data from 561 youth in Iceland (46% girls, M_{age} at Wave 1 = 14.3 years, SD = 0.3), only the multidimensional models (i.e., the three-factor and bifactor models) gave a good fit to the data. We then assessed the predictive power of the multidimensional models for academic achievement. The addition of academic achievement as an outcome variable to the bifactor model revealed that general school engagement, as well as specific behavioral engagement, predicted achievement. These findings are distinct from previous results using three-factor models, which indicated that behavioral engagement alone predicted later achievement. The results of the current study support the use of a bifactor model when using measures of school engagement.

Keywords

academic achievement, adolescence, bifactor models, school engagement

School engagement has been identified by researchers, teachers, and policy makers as an important asset for promoting school success and for addressing academic problems, such as school dropout and poor academic achievement (Appleton, Christenson, & Furlong, 2008; Fredricks, Blumenfeld, & Paris, 2004). School engagement[1] has been defined as '[the] student's active participation in academic and co-curricular or school related activities, and commitment to educational goals and learning. . . . It is a multidimensional construct that consists of behavioral (including academic), cognitive, and affective subtypes' (Christenson, Reschly, & Wylie, 2012, pp. 816–817).

The three aspects of school engagement, that is, students' cognitions, feelings, and behaviors, have been shown to be important for school success. Behavioral engagement reflects a student's willingness to participate in school related activities, such as attending classes. Emotional engagement includes a student's feelings about his or her school, such as his or her sense of belonging to school. Cognitive engagement describes a student's willingness to invest in cognitive abilities that relate to learning, such as self-regulated learning (Christenson et al., 2012; Li & Lerner, 2013).

The high correlations frequently observed among behavioral, emotional, and cognitive engagement (see e.g. Li & Lerner, 2013) has raised questions about the multidimensionality of school engagement. Acknowledging the extent to which a measure is multidimensional is important, as secondary dimensions can be lost when inappropriate models, such as one-factor models, or models

that do not acknowledge the common variance of the factors, are fit to multidimensional data (Ackerman, 1992; Reise, 2012). The current study examined the validity of the school engagement construct by comparing the model fit of three measurement models of school engagement (a one-factor, a three-factor and a bifactor model of school engagement) and by comparing the criterion validity of the good-fitting models by testing how strongly each predicts academic achievement.

The role of school engagement in academic success

In one of the earliest reviews of the concept of school engagement, Mosher and McGowan (1985) explained that physical presence in schools can be legislated (i.e., by making school attendance

---

[1] University of Iceland, Reykjavik, Iceland
[2] Oregon State University, USA
[3] Icelandic Education Testing Institute, Reykjavik, Iceland
[4] Tufts University, USA

Corresponding author:
Kristjan K. Stefansson, University of Iceland, Saemundargata 2, 101 Reykjavik, Iceland.
Email: kristjan@hi.is

mandatory), whereas school engagement cannot. A student's presence at school is necessary for, but does not guarantee, school engagement and the successful completion of compulsory school (Reschly & Christenson, 2012). For instance, Finn's (1989) Participation-Identification Model highlighted that successful school completion results from a gradual process of school engagement that involves a student identifying school-related goals and actively participating in school-related activities. As such, school engagement develops and entails bidirectional relations between students and their school contexts (Finn, 1989).

A relational developmental systems (RDS) perspective underlines this bidirectional nature of school engagement (Overton, 2015). In one of the most influential RDS models of Positive Youth Development (PYD), Lerner and colleagues (see e.g. Lerner, Phelps, Forman, & Bowers, 2009; Lerner, Lerner, & Benson, 2011) have identified school engagement as one of the key strengths of adolescents that, when aligned with ecological assets, promotes positive youth development. Similarly, influential researchers of motivation and school success, such as Eccles (2004), emphasize that the context, that is, schools, need to align better with the developmental needs of their students to support person-context fit and ensure that all students are motivated and engaged in their education.

There is considerable empirical support for the importance of school engagement for later school success (Reschly & Christenson, 2012). For example, a recent study found that 60% of high school dropouts could be identified based on their sixth-grade school engagement alone (i.e., attendance, misbehavior, and course failures; Balfanz, Herzog, & Mac Iver, 2007). Accordingly, one of the more important impacts of research on school engagement has been a shift away from focusing on individual characteristics, such as IQ, for promoting school success, and focus on understanding person-context fit in educational practices (Appleton et al., 2008; Fredricks et al., 2004).

The three-dimensional nature of school engagement

As already explained, the strong correlations among the subdimensions of school engagement have raised questions about the uniqueness of each subdimension (Fredricks et al., 2004). An example of the strong correlation among the subdimensions is a recent study on the factorial invariance of the Student Engagement Instrument (SEI) where subfactors of cognitive and emotional engagement were very strongly correlated among adolescences in Grades 6 through 12 ($r = .79$; Reschly, Betts, & Appleton, 2014).

The conceptual definition of school engagement has raised questions about the homogeneity of each subdimension (Betts, 2012). The heterogeneity of the subdimensions is reflected in their overlap with several other processes that have been identified by psychological and educational research, such as goal setting, self-regulated learning, social development, internal motivation, and reward contingencies (Betts, 2012). The overlap between individual subdimensions of school engagement and other processes has, for example, been observed in the strong correlation ($r = .51$ –.72) between measures of cognitive engagement and measures of motivational engagement among adolescences in Grades 9 through 12 (Betts, Appleton, Reschly, Christenson, & Huebner, 2010).

The wide range of concepts that relate to each subdimension have made it difficult to define three separate measures of engagement that include all the relevant aspects of each subdimension. Researchers have therefore frequently identified specific aspects of each subdimension and then used these attributes as indicators of general school engagement. This practice assumes that school engagement represents a unified, yet multidimensional, construct (Betts, 2012). As such, the bifactor model has been recommended as a potentially useful approach to partitioning the item variance into separate general and specific factors that can then be evaluated to better understand the structure of school engagement (Betts, 2012; Reise, 2012).

The bifactor measurement model

The bifactor measurement model, first described by Holzinger and Swineford (1937), has recently been rediscovered as an important approach to representing multidimensional measures in factor analysis and structural equation modeling (Reise, 2012; von Eye, Martel, Lerner, Lerner, & Bowers, 2011). The use of bifactor models allows researchers to examine a single common factor that represents a multidimensional construct, while also acknowledging the uniqueness of the individual dimensions that comprise it. More specifically, the bifactor model specifies that the covariance among a set of items can be accounted for by two processes; a single general factor that reflects the common variance among all the items, and a set of specific factors that reflect additional covariation among subsets of items. All factors are typically assumed to be orthogonal (i.e., uncorrelated), meaning that items representing different dimensions are hypothesized to correlate only because of their shared variance with the general factor (Betts, 2012; Reise, 2012; Reise, Morizot, & Hays, 2007).

To our knowledge, no study has yet measured school engagement as a single common construct while also recognizing its tripartite nature by using a bifactor model. The scarcity of such research was confirmed by searching for the words 'bifactor/bi-factor' and 'school/student engagement' in titles, keywords, and abstracts on the Web of Science™. This search revealed only three journal articles and one book chapter containing both words, none of which assessed all three dimensions of school engagement as defined by Christenson et al. (2012).

The current study

The current study is a four-wave longitudinal study that took place at the beginning and end of Grade 9 (Waves 1 and 2, respectively), and the beginning and end of Grade 10 (Waves 3 and 4, respectively) in Iceland. In the study, we examined the validity of a school engagement measure at Wave 1 in two ways, first, by comparing the fit of three models of school engagement and, second, by comparing the predictive validity of good-fitting models at Wave 1 for predicting scores on a standardized achievement test (the Icelandic National Examinations; INE) at Wave 3. Finally, we confirmed the reliability of the best fitting model by examining the longitudinal factorial invariance of the best fitting measure across all four waves.

More specifically, using the Behavioral-Emotional-Cognitive School Engagement Scale (BEC-SES; Li & Lerner, 2013), we tested three rival measurement models: a single-factor model of general school engagement; a three-factor model of behavioral, emotional, and cognitive school engagement; and bifactor model with school engagement as a general factor and three specific

behavioral, emotional, and cognitive engagement factors. Because previous research has found that the subdimensions of school engagement are highly correlated ($r > .50$; Li & Lerner, 2013), we followed the recommendation of Reise et al. (2007) and hypothesized that the measure of school engagement reflected two distinct sets of processes. First, we argued that students would exhibit systematic differences in how they would respond to all items, indicating a global school engagement factor. In addition, we hypothesized that each facet of school engagement (i.e., behavioral, emotional, and cognitive) would display systematic between-person variation that would be independent of the global school engagement factor. We therefore hypothesized that a bifactor model would fit our data better than either competing model.

After testing the rival measurement models, INE scores were added as an outcome variable to models that displayed good model fit. Based on previous research on school engagement and academic achievement (Chase, Hilliard, Geldhof, Warren, & Lerner, 2014; Christenson et al., 2012), we hypothesized that a three-factor model of school engagement would indicate a strong positive relationship between behavioral engagement and academic achievement, and a weak relationship between academic achievement and emotional and cognitive engagement. Based on this same research, we hypothesized that a general factor of a bifactor model of school engagement would positively predict later academic achievement. Due to the scarcity of research, the analyses on the relation of specific school engagement factors and academic achievement in the bifactor model were purely exploratory. Furthermore, based on previous research on the distortion that may occur when fitting inappropriate models to multidimensional data (Ackerman, 1992; Reise, 2012), we hypothesized that a bifactor model of school engagement would fit the data significantly better than a three-factor model, which would in turn fit the data significantly better than a unidimensional model. After selecting the best-fitting model, we tested factorial invariance of the best-fitting model across the four waves. We hypothesized that a configural, weak, and strong factorial invariance could be established across the four waves of measurement.

Method

The current study is part of a four-wave longitudinal investigation of adolescent development in Iceland conducted at the beginning of Grade 9 and lasting through the end of Grade 10 in the Icelandic compulsory school system. Academic achievement data was collected concurrently with the third wave of measurement and merged with the overall dataset.

Participants

We randomly selected 20 out of the 54 medium- to large-sized schools (>20 Grade 9 students) located in the Reykjavík capital area and the adjacent Reykjanes peninsula. Out of the 20 selected schools, 15 agreed to participate. Each participating school received a book as a gift for their school library for their participation. In order to increase the number of schools that participated, and thereby ensuring more diverse responses at the school level, two classrooms in each school were selected at random in schools that had more than two classrooms. These 30 classrooms had a total of 625 students. A total of 561 parents (90%) gave written consent for their child's participation, and 539 (96%) of youth with parental

Table 1. Sample descriptive statistics at Wave 1.

Females n (%)	249 (46.20)
Age mean (SD)	14.28 (0.28)
Mothers with only compulsory education n (%)	77 (14.81)
Foreign language spoken at home* n (%)	31 (5.82)
Fewer than 25 books at home n (%)	65 (12.22)
Behavioral-Emotional-Cognitive School Engagement Scale (1–4) mean (SD)	3.32 (0.45)
Behavioral engagement (1–4) mean (SD)	3.49 (0.45)
Emotional engagement (1–4) mean (SD)	3.04 (0.56)
Cognitive engagement (1–4) mean (SD)	3.41 (0.56)

Note. Total number of participants = 561. *In Iceland, a foreign language spoken at home frequently serves as an indicator of a household minority status.

consent participated at Wave 1 (mean age 14.3, $SD = 0.3$, 46% girls). The population in Reykjavík and the Reykjanes area includes 66.4% of all Icelandic children and is socially heterogeneous (Table 1).

Procedure

Participants completed a paper-and-pencil survey during a 40-minute school visit by trained research staff. Standardized instructions were used to ensure that data collection was uniformly administered. Students who were absent during the school visit were contacted by e-mail, mail, or phone, and asked to complete and return the survey by mail.

Measures

We describe our measures below. For each measure, the model-based reliability estimate coefficient ω (McDonald, 1999) was calculated to indicate the proportion of the scale variance that was due to all common factors (Zinbarg, Revelle, Yovel, & Li, 2005). Coefficient ω is analogous to coefficient α (Reise, 2012); therefore reliability estimates above the .70 level were interpreted as indicators of adequate reliability (Kline, 2011).

Behavioral-Emotional-Cognitive School Engagement Scale (BEC-SES). To measure school engagement, we used the Behavioral-Emotional-Cognitive School Engagement Scale (BEC-SES) developed by Li and Lerner (2013). The BEC-SES consists of the three subscales of school engagement: behavioral, emotional, and cognitive. Each subscale was measured using five items administered using a four-point Likert scale (answer options differed across scales, see below). Abbreviated item content can be seen in Table 2.

A three-factor model of BEC-SES has shown evidence of strong measurement equivalence between boys and girls, between youth of different socioeconomic status, and across youth in US Grades 9 through 11 (Li & Lerner, 2012). The measure was translated into Icelandic by two independent translators. The translations were reconciled by researchers fluent in both languages and pretested with 77 Grade-9 students from a single school. Coefficient ω for the whole BEC-SES in the current sample was .95.

Behavioral engagement. Behavioral engagement included five items whose content ranged from shallow engagement (e.g., class attendance) to deep engagement (e.g., effort). The subscale focuses on students' voluntary behaviors within the school context to minimize possible confounding effects of non-student related variables

Table 2. The Behavioral-Emotional-Cognitive School Engagement Scale. Abbreviated item content and frequencies at Wave 1.

No.	S	Abbreviated item content	C1%	C2%	C3%	C4%\
1	Behavioral engagement	Come to class unprepared*	3.2%	22.4%	47.4%	26.9%
2	Behavioral engagement	Complete homework on time	0.9%	7.4%	16.2%	75.6%
3	Behavioral engagement	Skip classes without permission*	0.2%	2.9%	10.1%	86.9%
4	Behavioral engagement	Take part in group (class) discussions	2.7%	12.6%	33.0%	51.7%
5	Behavioral engagement	Work hard to do well in school	0.9%	5.0%	26.2%	67.9%
6	Emotional engagement	Feel part of my school	6.6%	10.4%	51.3%	31.6%
7	Emotional engagement	Care about the school	7.0%	14.2%	46.3%	32.5%
8	Emotional engagement	Happy to be at my school	3.8%	10.1%	37.7%	48.5%
9	Emotional engagement	Don't find school fun and exciting*	9.7%	21.0%	41.7%	27.6%
10	Emotional engagement	Enjoy the classes I am taking	5.2%	15.8%	58.7%	20.3%
11	Cognitive engagement	Learn as much as I can at school	2.7%	7.5%	41.1%	48.7%
12	Cognitive engagement	Is important to make good grades	2.0%	3.6%	20.5%	74.0%
13	Cognitive engagement	The things I learn at school are useful	2.5%	11.3%	46.1%	40.0%
14	Cognitive engagement	Think a lot about how to do well in school	3.8%	17.1%	42.0%	37.2%
15	Cognitive engagement	School is very important for later success	1.4%	2.7%	22.3%	73.6%

Note. These are average results over 20 data sets. *Reverse-worded item.

(an example of academic behaviors outside the school context that can be confounded by non-student variables, is participation in private tutoring, which may be related to social economic status). For each item, respondents were asked to rate how often they engaged in specific behaviors using a scale from 1 (*never*) to 4 (*always*). Coefficient ω for behavioral engagement in the current sample was .82.

Emotional engagement. The emotional engagement subscale included five items that assessed students' sense of belonging and their affect toward school. Happiness, excitement, and enjoyment were used to measure three related, yet distinct, types of positive affect. Items to tap school connectedness assessed different aspects of the emotional relationships students had with their school and classes. The respondents were asked indicate their agreement to five emotional statements on a scale from 1 (*strongly disagree*) to 4 (*strongly agree*). Coefficient ω for emotional engagement in the current sample was .87.

Cognitive engagement. Cognitive engagement was measured by five items designed to assess the extent to which students valued education and things learned at school, as well as their thoughts about learning. More specifically, goal orientation, identification with school, and perceptions of the link between students' lives and school were included as core indicators of cognitive engagement. The respondents were asked indicate their rate of agreement to five cognitive statements on a scale from 1 (*strongly disagree*) to 4 (*strongly agree*). Coefficient ω for cognitive engagement in the current sample was .90.

Academic achievement. Icelandic National Examinations (INE) scores were retrieved from the Icelandic Educational Testing Institute (IETI; 2014). The IETI administers an annual INE in language skills (Icelandic), mathematics and English at the beginning of Grade 10. The exam is multidimensional and includes subcomponents that measure, for example, algebra, geometry, grammar, and spelling. The standardized scores range from 0 to 60 with a mean of 30 and a standard deviation of 10. The single academic achievement factor was fit to the three observed test scores using the direct maximum likelihood estimator. A unidimensional model was saturated, $\chi2$ (0) = 0, $p < .001$; CFI = 1.00; RMSEA = .00, 90% CI (.00, .00). The academic achievement factor was clearly manifested in

the total test scores of Icelandic, mathematics, and English with standardized factor loadings of 0.93, 0.83, and 0.72, respectively. Coefficient ω for academic achievement in the current sample was .88.

Data analysis

A series of factor analyses and structural equation models was estimated using version 7.3 of the Mplus software package (Muthén & Muthén, 1998–2012). The estimates of latent factors were scaled using the fixed factor method (see Little, 2013), setting the variance of each latent factor to unity. For the BEC-SES, a bifactor model was defined where each specific factor was indicated by the items suggested by the previously established three-factor model (see Li & Lerner, 2012). In addition, we defined a global school engagement factor that was indicated by all the items across the three specific factors. No cross-loadings or item-correlations were allowed. In addition, for identification purposes of the bifactor model, the correlations between all latent factors (general and specific) were set to zero within and across measurements.

Model fit was estimated by evaluating several fit indices: the chi-square statistic for the WLSMV estimation method, the root mean square error of approximation (RMSEA), and the comparative fit index (CFI). Smaller chi-square and RMSEA values (RMSEA $\leq$ .06), and higher CFI values ($\geq$ .95) indicate a good model fit (West, Taylor, & Wu, 2013). Differences in model fit were confirmed with a chi-square difference tests using the DIFFTEST option in Mplus. The chi-square difference tests were further supplemented by comparing Bayesian information criterion (BIC) values, where a smaller value indicates a better fit.

Due to the low number of response options for our Likert-type data, we treated all indicators as categorical and estimated all models using robust weighted least squares (WLSMV). During the four waves of measurement, 91%, 86%, 90%, and 87% of the participants had complete data on all the school engagement items, respectively. After the last wave, 68% of the participants had complete data across the four waves of measurement. We considered the missing data to be missing at random (MAR). Correlational analysis revealed significant correlations between several variables, such as self-reported grades, mother's education, father's education, and mother's occupation and missing cases at later waves. These background variables

were used to inform the creation of 20 imputed datasets without missing values using the multiple imputation feature of Mplus.

The results in the manuscript are, as noted, the pooled results from 20 imputations with one exception, this exception is the difference testing of nested models at Wave 1 using the DIFFTEST feature of Mplus. The DIFFTEST feature is currently not available for the analysis of imputed data. Given the small amount of missingness during Wave 1 (9%), and given that comparative analysis using imputed and non-imputed data showed very similar results, we based the nested model comparisons on non-imputed data. Because this approach uses a pairwise present approach to address missingness, we conditioned all items on the same covariates as used to inform the imputed datasets.

We calculated the intraclass correlation (ICC) for all the items used in the analysis for both class and school level in a series of two-level unconditional models. All ICC values were lower than .10, which has been considered a minimum to produce appreciable bias in standard errors if multilevel statistical techniques are not used (Kline, 2011). To minimize the risk of making a Type 1 error, we ran all the CFA and SEM models twice producing correct standard errors using a sandwich estimator, first based on the class level variation, and again based on the school level variation. The CFA models showed no appreciable bias in standard errors under either condition. The SEM models, however, showed an appreciable bias in standard errors for the regression coefficients when clustering on school level was not taken into account. We therefore decided, as we are only interested in the individual level, to use the COMPLEX feature of Mplus and produce correct standard errors using a sandwich estimator based on the school level clustering.

We established configural, weak, and strong longitudinal factorial invariance for the bifactor model using a method for models with ordered-categorical data described by Millsap and Yun-Tein (2004). Residual variances of same indicators at all waves were allowed to correlate. Furthermore, we allowed cross-time stability correlations among same factors, but no correlations were allowed with other factors, within or across measurement occasions (see Little, 2013). We evaluated the invariance constraints using a guideline made by Chung and Rensvold (2002), where a change of more than .01 in the comparative fit index (CFI) indicates that the assumption of invariance does not hold.

Results

Abbreviated item content and frequencies for the 15-item Behavioral-Emotional-Cognitive School Engagement Scale are summarized in Table 2. Items were not markedly skewed with the exception of Item 3, in response to which 87% of the participants said they had never skipped class without permission. The item was retained, as the WLSMV method has generally performed well with skewed ordered categorical variables when sample sizes are not small (about $N = 200$; Kline, 2011).

Confirmatory factor analyses of the BEC-SES

The WLSMV estimation method was used to fit three measurement models to the data: a one-factor model, a three-factor model, and a bifactor model. Model identification was established by fixing the variance of each latent variable to unity. Model fits are summarized in Table 3. The one-factor model exhibited inadequate fit, $\chi^2(90) = 552.37$; CFI = .90; RMSEA = .10, because of large chi-square and low CFI value. The three-factor model showed an acceptable fit, $\chi^2(87) = 227.68$; CFI = .97; RMSEA = .05, with a significant reduction in the chi-square value compared to the nested one-factor model. In addition, the model showed an acceptable RMSEA value and a good CFI value. The bifactor model, however, provided the best fit of the three models, with the lowest chi-square value and good RMSEA and CFI values, $\chi^2(75) = 149.89$; CFI = .98; RMSEA = .04. A chi-square difference test using the DIFFTEST option in Mplus confirmed that the three-factor model fit the data better than the nested one-factor model, $\Delta\chi^2(3) = 211.65$, $p < .001$, and that the bifactor model fit the data better than the three-factor model, $\Delta\chi^2(12) = 87.84$, $p < .001$. The DIFFTEST results were further supplemented by estimating the Bayesian information criterion (BIC). The BIC for the bifactor model was substantially lower than the BIC values for the two other models (see Table 3), indicating a better fit for the bifactor model (see Raferty, 1995).

The standardized factor loadings of the different models can be seen in Table 4. The one-factor model was well defined and highly reliable ($\omega = .93$), with factor loadings ranging from .42 (participation in classroom discussions) to .79 (caring about the school). The three-factor model was well defined and reliable (behavioral engagement, $\omega = .82$; emotional engagement, $\omega = .87$; and cognitive engagement, $\omega = .90$), with factor loadings ranging from 0.49 to 0.88 for the same items as the minimum and maximum factor loadings in the one-factor solution. The general school engagement scale in the bifactor model was also well defined and highly reliable ($\omega = .93$), with factor loadings ranging from 0.40 (come to class unprepared) to 0.76 (learn as much as I can at school). The difference in maximum and minimum factor loadings between models indicates that general school engagement has a qualitatively different meaning when each specific factor (i.e., behavioral, emotional, and cognitive) has been separated from the general school engagement factor. Although also highly reliable (behavioral engagement, $\omega = .84$; emotional engagement, $\omega = .87$, and cognitive engagement, $\omega = .91$), the three specific factors were less well defined than the general factor. All the specific factor loadings were significant at the $p < .01$ level, although one behavior engagement factor item, which refers to participation in class discussions (Item 4), showed a particularly low loading (0.11).[2]

All of the subfactors in the three-factor model correlated strongly with each other, with latent correlation coefficients ranging from $r = .65$ between emotional and cognitive engagement to $r = .72$ between cognitive and behavioral engagement. The remaining correlation between emotional and behavioral engagement was $r = .66$.

Table 3. CFA fit statistics for the Behavioral-Emotional-Cognitive School Engagement Scale measurement models at Wave 1.

	χ^2	s	$\Delta\chi^2$	Δdf	CFI	RMSEA	BIC
One-factor model	552.373	15.52	90			.90	.10	16324.44
Three-factor model	227.675	8.17	87	324.70	3	.97	.05	15808.16
Bifactor model	149.885	5.38	75	77.79	12	.98	.04	15772.19

Note. These are average results over 20 data sets. The BIC was retrieved with a separate CFA using maximum likelihood estimation; Chi-square difference tests between nested models at Wave 1 were conducted with non-imputed data.

Table 4. Standardized factor loadings of the three measurement models for the Behavioral-Emotional-Cognitive School Engagement Scale at Wave 1.

Item	One-factor School engagement scale	Three-factor			Bifactor			
		Behavioral engagement	Emotional engagement	Cognitive engagement	School engagement scale (general)	Behavioral engage-ment (specific)	Emotional engage-ment (specific)	Cognitive engage-ment (specific)
1	0.48**	0.56**			0.40**	0.82**		
2	0.70**	0.80**			0.67**	0.46**		
3	0.67**	0.74**			0.64**	0.30**		
4	0.42**	0.49**			0.44**	0.11**		
5	0.70**	0.85**			0.74**	0.20**		
6	0.74**		0.81**		0.64**		0.45**	
7	0.79**		0.88**		0.64**		0.67**	
8	0.75**		0.83**		0.62**		0.56**	
9	0.55**		0.62**		0.47**		0.41**	
10	0.58**		0.66**		0.56**		0.29**	
11	0.75**			0.81**	0.76**			0.21**
12	0.75**			0.85**	0.74**			0.39**
13	0.73**			0.76**	0.60**			0.51**
14	0.72**			0.77**	0.64**			0.47**
15	0.74**			0.80**	0.60**			0.67**

Note. These are average results over 20 data sets.
**$p < 0.01$.

Criterion validity: Latent regression analyses of school engagement and academic achievement

INE scores at Wave 3 were added—as a continuous outcome variable—to the three-factor and the bifactor measurement models from Wave 1 to assess the relative performance of the different measurement models in predicting academic achievement (see Figure 1). The one-factor model was not included due to the poor model fit established in the CFA.

The WLSMV estimation method was again used to fit these structural equation models to the data, and model identification was again enabled by setting the variance of each latent variable to unity. The fit indices of the models and latent regressions are shown in Table 5. The three-factor and the bifactor models showed a good fit; the bifactor model at Wave 1 fit the data significantly better than the three-factor model, according to a chi-square difference test using the DIFFTEST option in Mplus, $\Delta\chi^2(13) = 61.25$, $p < .001$.

In the three-factor model, only the behavioral engagement factor strongly predicted subsequent INE scores, $\beta = 0.73$, 95% CI (0.49, 0.98). In contrast, the bifactor model at Wave 1 produced two separate direct effects. The general school engagement factor produced a strong direct effect, $\beta = 0.51$, 95% CI (0.37, 0.65) and, in addition, the specific behavioral engagement factor produced a moderate direct effect, $\beta = 0.25$, 95% CI (0.06, 0.44) on the INE scores. The specific emotional and specific cognitive factors had weak and non-significant effects. The bifactor and three-factor school engagement models at Wave 1 both explained 36% of the variance of the INE scores.

Factorial invariance of the bifactor model of BEC-SES

Finally, in order to ensure that the structure of school engagement does not substantially vary over time, our last analytic step was to test factorial invariance of the bifactor solution, we examined the consistency of measurement of the bifactor model by establishing configural, weak, and strong factorial invariance across the four waves of available data. Scale identification was obtained by using guidelines described by Millsap and Yun-Tein (2004), the results can be seen in Table 6. The configural invariance model showed excellent fit with an average CFI of .978 and a standard deviation of only .001 across the 20 datasets. The weak invariance model was specified by fixing the individual factor loadings to be equal across the four waves. This specification caused a very small improvement in model fit, increasing the CFI by .001 while the standard deviation of the CFI remained small (.001). The strong invariance model was further specified by fixing individual thresholds to be equal across the four waves. The strong invariance model gave the same CFI and standard deviation as the weak invariance model. Differences in CFI between invariance models were well below the .01 criterion chosen for the comparison, which supported configural, weak, and strong factorial invariance across the four waves.

Discussion

This study contributes to the growing school engagement literature in three ways as will be explained in the following sections. To summarize, first, the study confirms the tripartite nature of school engagement. Second, the significantly better fit of the bifactor model suggests that, rather than being strictly unidimensional or adhering strictly to a tripartite structure, the construct of school engagement may be conceptualized as having multiple dimensions that share substantial overlap. The bifactor model confirmed that a reliable general school engagement factor underlies all of the school engagement items of the BEC-SES, regardless of their behavioral, emotional, and cognitive origin. Third, the bifactor model showed that the school engagement items gave rise to three specific factors.

Furthermore, the diverse factor loadings of the specific factors indicated that behavioral, emotional, and cognitive engagement, above and beyond general school engagement, was poorly defined in the BEC-SES. The poor definition of the specific factors can be
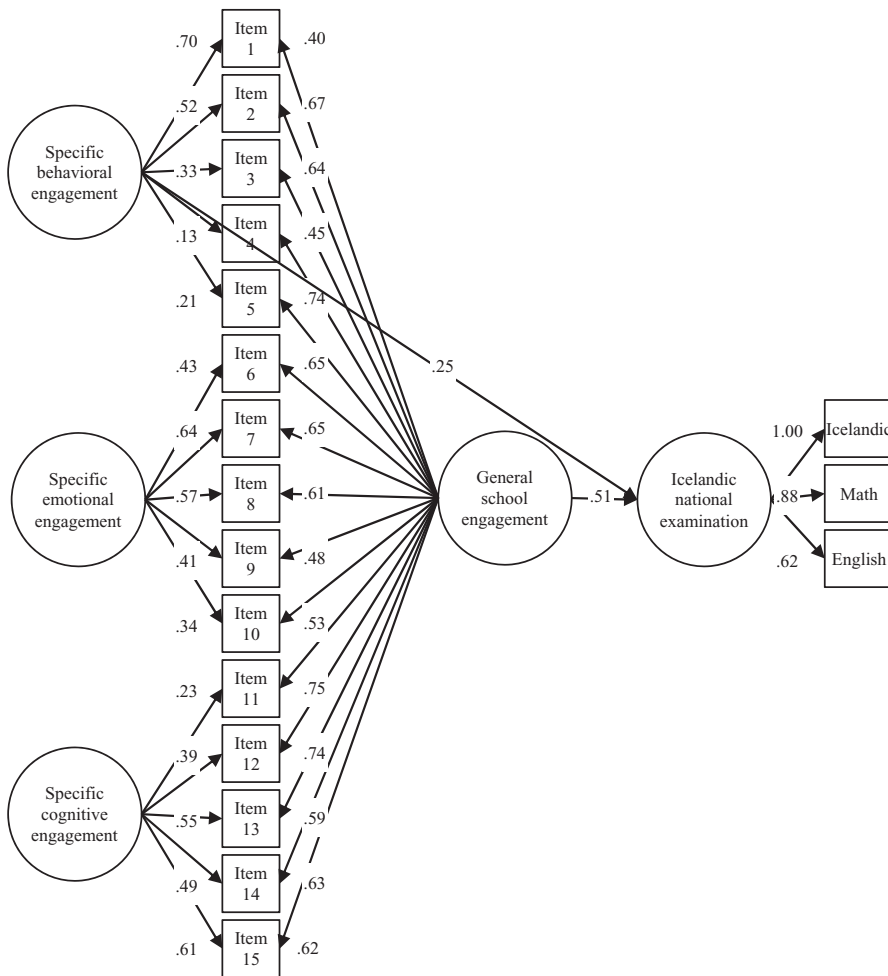
Figure 1. Empirical results of a structural equation model where a bifactor model of school engagement at the beginning of Grade 9 (Wave 1) predicts Icelandic national examination scores at the beginning of Grade 10 (Wave 3). Total number of participants = 561. The variances of the latent factors were set to unity to allow for identification. For clarity, only significant ($p < .01$) factor loadings and regression coefficients are shown in the diagram. Non-significant regression coefficients and fit indices can be found in Table 5.

used to inform further development in the measurement of school engagement. The specific emotional factor can, for example, be refined by separating the factor into two specific emotional engagement factors, one that it sensitive to the emotional engagement in the school in general, and another that is sensitive to emotional engagement in classes specifically.

Comparing rival models of school engagement

The three-factor and bifactor models gave adequate and good fit to the data, respectively. The one-factor model, however, fit the data poorly. The poor fit of the one-factor model provides a reason not to encourage the use of a one-factor model with the BEC-SES items. The three-factor and bifactor models both accounted for the multidimensional nature of school engagement, giving both models a much better fit than the one-factor model. However, the high correlation among the subscales in the three-factor model became problematic when the three-factor model was used to predict INE scores, as the majority of the effect was due to a general school engagement factor, which was not modeled in the three-factor model. Because of the lack of the general school engagement dimension in the three-factor model, we do not recommend it as a means to represent school engagement. The bifactor model, on the other hand, took the multidimensional nature of school engagement into account by modeling a general factor and specific subfactors, in essence combining the strengths of the one- and three-factor models.

In addition, the bifactor model provided valuable information about the meaning of the school engagement construct. The general

Table 5. SEM fit statistics and standardized regression coefficients of Behavioral-Emotional-Cognitive School Engagement Scale models at Wave 1 predicting Icelandic National Examinations scores (Icelandic, mathematics, and English combined) at Wave 3.

	χ^2	s	df	CFI	RMSEA	R^2	β
Three-factor model	263.462	7.09	129	.97	.04	.36	
Behavioral engagement							0.73**
Emotional engagement							−0.09
Cognitive engagement							−0.12
Bifactor model	204.369	5.21	116	.98	.04	.36	
General school engagement							0.51**
Specific behavioral engagement							0.25**
Specific emotional engagement							−.12
Specific cognitive engagement							−.13

Note. These are average results over 20 data sets; Chi-square difference tests between nested models at Wave 1 were conducted with non-imputed data.
**$p < .01$; *$p < .05$.

factor had adequate loadings from all items, suggesting that all of the items were adequate manifestations of general school engagement. When looking at each of the subcomponents, the specific behavioral factor loadings were very heterogeneous. The specific behavior engagement factor was defined by strong factor loadings for Item 1 (coming to class unprepared) and Item 2 (completing homework on time). These strong factor loadings indicated that the specific behavioral factor modeled in this study was mainly a measure of academic behavior. In contrast, for example, Item 5 (work hard to do well in school) had a very weak factor loading for the specific behavioral factor but a very strong factor loading for the general school engagement factor. The low factor loading indicates that working hard to do well at school is a poor manifestation of the current definition of specific behavioral engagement but, instead, represents general school engagement very well.

The specific emotional engagement factor also had heterogeneous factor loadings, with the strongest factor loading for an item that assessed how much students cared about their schools. The item with the weakest factor loading for the specific emotional factor (Item 10) differed conceptually from the other emotional items. Item 10 referred to the classes the student was taking and did not refer to the school in general as the other items. This finding suggests that emotional engagement in classes may be a different specific dimension than more general emotional engagement in the school and, as such, general school engagement may be better represented by separating the specific emotional engagement dimension by different contexts of the school environment.

The specific cognitive engagement factor also had heterogeneous factor loadings, with the strongest factor loadings for items that index the importance of school for later success (Item 15) and the usefulness of things learned in school (Item 13). This finding indicates that the specific cognitive engagement modeled in this study was mainly a measure of the practical importance of school. The weakest factor loading for the specific cognitive engagement factor was associated with Item 11 (learn as much as I can). Item 11 differs from the other cognitive items as it closely relates to internal motivation. The low factor loading indicates that internal motivation is a poor manifestation of the current definition of specific cognitive engagement, but that it is a strong indicator of global engagement.

Taken together, the factor loadings of the specific factors of the bifactor model demonstrate that if the specific factors are to be used

as valid measures to inform practice, each of the three specific factors should be further separated into more fine-grained and better defined specific factors. In contrast, the factor loadings of the general factor show that the general factor is strongly manifested in all the proposed items and fits the data very well when the specific factors have been parsed out.

Predicting academic achievement with different models of school engagement

As hypothesized, the three-factor and the bifactor models of school engagement at Wave 1 both positively predicted academic achievement at Wave 3. The three-factor model strongly predicted academic achievement, but only through the behavioral factor. The bifactor model strongly predicted academic achievement through the general school engagement factor. In addition, the bifactor model predicted academic achievement through a specific behavioral factor, but only to a moderate degree over and above the general school engagement factor. As the specific behavioral factor was, by definition, not related to general school engagement, the measure represents students who chose to engage, or not to engage, academically in school-related activities regardless of the student's level of general school engagement. A student with high specific behavioral engagement could therefore attend school and do his or her homework without necessarily caring about school or finding school important. The results indicated that such behaviors positively predicted academic achievement to a moderate degree. However, the results also indicated that general school engagement, when behavioral, emotional, and cognitive aspects are included, is a separate factor that strongly predicted academic achievement.

The findings from the latent regression analyses suggest that interventions aimed at improving academic achievement may have a considerable effect if they focus on all aspects (i.e., behavioral, emotional, and cognitive) of school engagement. This is in sharp contrast to previous findings that indicate that, of the three school engagement dimensions, only behavioral engagement predicts academic achievement (see Chase et al., 2014). In addition, interventions that include an additional emphasis on aspects of specific behavioral engagement (i.e., feeling prepared for class, finishing homework, attending class) are likely to be associated with additional improvement in academic achievement. Further research is needed to develop and confirm the existence of the specific behavioral engagement factor and its implications for academic achievement.

Limitations, strengths and implications for future research

Some issues should be considered when interpreting the results of this study. A key limitation is that the major source of information was self-report. This form of data collection may bias the results, as the observed correlations between the different items may be due to common method variance rather than representing actual relations among underlying latent constructs. An exception to the self-report data-collection, and a strength of the study, was the inclusion of INE scores, which provided a more valid and normally distributed measure of academic achievement than the frequently used self-reported grades. In general, research would benefit from a cross-

Table 6. Model fit statistics for the tests of measurement invariance of general and specific aspects of behavioral, emotional and cognitive engagement across four waves.

	χ^2	s	df	$\Delta\chi^2$	Δdf	CFI	s	RMSEA	s
Configural model	1843.028	9.43	1536			.978	.001	.019	.000
Weak invariance	1909.856	11.25	1614	66.83	78	.979	.001	.018	.000
Strong invariance	1987.091	11.49	1692	77.24	78	.979	.001	.018	.000

Note. These are average results over 20 data sets.

validation of the BEC-SES obtained from additional sources, such as from parent and teacher reports or through classroom observations.

We wish to emphasize, when discussing the findings of this study, that the predictive effects do not imply causation. However, the predictive effects in this study highlighted the consequences of mis-specifying multidimensional models with highly correlated subfactors. Future research into the factors associated with, and development of school engagement should consider multiple covariates and methods to better understand individual and group changes during adolescence. Another limitation to this study is that it was conducted with a limited age range in a homogeneous cultural area. This sampling restricts the generalizability of the research results to students of different ages and the results may not represent the findings based on youth from other cultures or subgroups within the Icelandic population. The research results would benefit from a cross-cultural and cross-group validation of future studies.

In general however, the results confirm that school engagement 'consists of behavioral (including academic), cognitive, and affective subtypes' as defined by Christenson and colleagues (2012, pp. 816–817), without any one subtype outweighing the other two. The significantly better fit of the bifactor model suggests that, rather than being unidimensional or multidimensional, school engagement is characterized by both a single and multiple dimensions. Furthermore, the results showed that important secondary dimensions can be lost when using nested models, such as the three-factor model. Our results suggest that a bifactor model is the best way to represent a comprehensive measure of school engagement.

Funding

Notes

1. In the research literature, the terms school engagement and student engagement are often used interchangeable (Libbey, 2004), however, no distinction is made between these two terms in this article.
2. As recommended by Reise (2012), an exploratory bifactor analysis was conducted prior to the CFA of the bifactor structure. The EFA was conducted using three- to five-factor solutions on all available waves. Item 4 cross-loaded on several factors but only a number of the factors had substantive meanings. The most consistent cross-loadings were on the emotional engagement factor using a four-factor solution. Item 4 had a .08, .11, .31, and .29 cross-loading on specific emotional engagement at Wave 1 through Wave 4, respectively. These cross-loadings were similar to the loadings on the theoretically assigned factor (specific

behavioral engagement) .14, .16, .13, and .26, respectively. We decided not to remove the item from its theoretically assigned factor, and not to include the cross-loadings in the CFA, but alert researchers of the risk of parameter distorting effects due to cross-loadings on group factors.

References

Ackerman, T. A. (1992). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. *Journal of Educational Measurement*, *29*(1), 67–91. doi:10.1111/j.1745-3984.1992.tb00368.x

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, *45*(5), 369–386. doi:10.1002/Pits.20303

Balfanz, R., Herzog, L., & Mac Iver, D. J. (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. *Educational Psychologist*, *42*(4), 223–235.

Betts, J. (2012). Issues and methods in the measurement of student engagement: Advancing the construct through statistical modeling. In S. L. Christenson, L. A. Reschly & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 783–805). New York, NY: Springer.

Betts, J. E., Appleton, J. J., Reschly, A. L., Christenson, S. L., & Huebner, E. S. (2010). A study of the factorial invariance of the Student Engagement Instrument (SEI): Results from middle and high school students. *School Psychology Quarterly*, *25*(2), 84–93. doi:10.1037/A0020259

Chase, P. A., Hilliard, L. J., Geldhof, G. J., Warren, D. J. A., & Lerner, R. M. (2014). Academic achievement in the high school years: The changing role of school engagement. *Journal of Youth and Adolescence*, *43*(6), 884–896. doi:10.1007/s10964-013-0085-4

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. doi:10.1207/S15328007sem0902_5

Christenson, S. L., Reschly, L. A., & Wylie, C. (Eds.). (2012). *Handbook of research on student engagement*. New York, NY: Springer.

Eccles, J. S. (2004). Schools, academic motivation, and stage-environment fit. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of Adolescent Psychology* (2nd ed.). Hoboken, NJ: Wiley.

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research*, *59*(2), 117–142. doi:10.3102/00346543059002117

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109.

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychmetrika*, *2*(1), 41–54.

Icelandic Educational Testing Institute. (2014). The homepage of the Icelandic Educational Testing Institute. Retrieved from http://www.namsmat.is

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.

Lerner, J. V., Phelps, E., Forman, Y., & Bowers, E. P. (2009). Positive youth development. In R. M. Lerner & L. D. Steinberg (Eds.), *Handbook of adolescent psychology* (3rd ed., *Vol. 1*, pp. 524–558). Hoboken, NJ: Wiley.

Lerner, R. M., Lerner, J. V., & Benson, J. B. (2011). Positive youth development: Research and applications for promoting thriving in adolescence. *Advances in Child Development and Behavior*, *41*, 1–17.

Li, Y. B., & Lerner, R. M. (2012). *Testing across-group and longitudinal measurement equivalence of a tripartite measure of adolescent school engagement*. Washington, DC: American Institutes for Research.

Li, Y. B., & Lerner, R. M. (2013). Interrelations of behavioral, emotional, and cognitive school engagement in high school students. *Journal of Youth and Adolescence*, *42*(1), 20–32. doi:10.1007/s10964-012-9857-5

Libbey, H. P. (2004). Measuring student relationships to school: Attachment, bonding, connectedness, and engagement. *Journal of School Health*, *74*(7), 274–283. doi:10.1111/j.1746-1561.2004.tb08284.x

Little, T. D. (2013). *Longitudinal structural equation modeling*. New York: Guilford Press.

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

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research*, *39*(3), 479–515. doi:10.1207/S15327906mbr3903_4

Mosher, R., & McGowan. (1985). *Assessing student engagement in secondary schools: Alternative conceptions, strategies of assessing, and instruments*. University of Wisconsin: Research and Development Center.

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus User's Guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Overton, W. F. (2015). Processes, relations, and relational-developmental-systems. In R. M. Lerner, W. F. Overton, & P. C. M. Molenaar (Eds.), *Handbook of child psychology and developmental science* (7th ed., *Vol. 1*, pp. 9–63). Hoboken, NJ: Wiley.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163. doi:10.2307/271063

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*(5), 667–696. doi:10.1080/00273171.2012.715555

Reise, S. P., Morizot, J., & Hays, R. D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research*, *16*, 19–31. doi:10.1007/s11136-007-9183-7

Reschly, A. L., Betts, J., & Appleton, J. J. (2014). An examination of the validity of two measures of student engagement. *International Journal of School & Educational Psychology*, *2*(2), 106–114. doi:10.1080/21683603.2013.876950

Reschly, A. L., & Christenson, S. L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. *Handbook of Research on Student Engagement*, 3–19. New York, NY: Springer.

von Eye, A., Martel, M. M., Lerner, R. M., Lerner, J. V., & Bowers, E. P. (2011). Interpreting theory and method in the study of positive youth development: The sample case of gender-specificity and longitudinal stability of the dimensions of intentional self-regulation (selection, optimization, and compensation). In R. M. Lerner, J. V. Lerner & J. B. Benson (Eds.), *Positive youth development. Advances in child development and behavior* (Vol. *41*, pp. 347–372). London, UK: Elsevier.

West, S., Taylor, A., & Wu, W. (2013). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). NY: Guilford.

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's alpha, Revelle's beta, and McDonald's (omega H): Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, *70*(1), 123–133. doi:10.1007/s11336-003-0974-7

Paper III: Stefansson, K. K., Gestsdottir, S. Birgisdottir, F. Lerner, R. M. (manuscript submitted). School engagement and intentional self-regulation: A reciprocal relation in adolescence.

School engagement and intentional self-regulation: A reciprocal relation in adolescence

Kristjan K. Stefansson, Steinunn Gestsdottir, and Freyja Birgisdottir

University of Iceland

Richard M. Lerner

Tufts University

Kristjan K. Stefansson is a Ph.D. student at the Faculty of Education Studies, University of Iceland. Dr. Steinunn Gestsdottir is a Professor at the Faculty of Psychology, University of Iceland. Dr. Freyja Birgisdottir is an Associate Professor at the Faculty of Psychology, University of Iceland. Dr. Richard M. Lerner is the Bergstrom Chair in Applied Developmental Science and the Director of the Institute for Applied Research in Youth Development at Tufts University.

Correspondence concerning this article should be addressed to Kristjan K. Stefansson, Faculty of Education Studies, University of Iceland, Saemundargata 2, 101 Reykjavik, Iceland. E-mail: kristjan@hi.is

Abstract

Students' gradual disconnection from school in adolescence, as reflected in decreased school engagement, has been demonstrated in several cultures and is of great concern to educators. At the same time, intentional self-regulation (ISR) has been shown to be a precursor, mediator, and outcome of school engagement. However, the relation between school engagement and ISR during adolescence is poorly understood. In this research, we explored the reciprocal relation between school engagement and ISR during adolescence. Based on a sample of 561 adolescents in Iceland (46% girls; Mage at Wave 1 = 14.3 years; $SD = 0.3$) and four waves of data collected at the beginning and end of Grades 9 and 10, the results demonstrated a reciprocal relation between school engagement and ISR after controlling for several covariates (e.g., gender and academic achievement). Furthermore, the results indicated decreased stability of both school engagement and ISR during the observed period. A decreasing stability is consistent with theories that present school engagement and ISR as malleable constructs that are open to contextual influences. Implications of the findings for future research are discussed.

Keywords: School engagement, student engagement, intentional self-regulation, adolescence, reciprocal relations.

School engagement and intentional self-regulation: A reciprocal relation in adolescence

The importance of school engagement for school success, such as good academic achievement and low dropout rates, has been well established (see e.g., Christenson, Reschly, & Wylie, 2012; Tuominen-Soini & Salmela-Aro, 2014). At the same time, recent studies have shown that half of American students are not sufficiently engaged in school (e.g., Gallup Student Poll, 2015), which correspond to findings from studies across the world (OECD, 2012a). Importantly, school engagement also has been considered to be malleable and, as such, open to contextual influences, including influences from parents, teachers, as well as students themselves (Appleton, Christenson, & Furlong, 2008; Fredricks, Blumenfeld, & Paris, 2004). Accordingly, researchers and educators have called for research that seeks to identify the variables promoting school engagement in the classroom (Coalition for Psychology in Schools and Education, 2006; Shernoff, 2013).

Intentional self-regulation (ISR) has been suggested to be an important precursor, mediator, and outcome of school engagement (Cleary & Zimmerman, 2012; Wolters & Taylor, 2012). However, there is limited empirical support for the proposed reciprocal relation between school engagement and ISR and research on the two constructs has suffered from conceptual confusion and measurement issues. Consequently, scholars have called for a better understanding on the degree of overlap between school engagement and ISR skills (Boekaerts, 2016; Eccles, 2016). The current study takes a step towards such an understanding by assessing potential reciprocal relations between school engagement and ISR among adolescents in Iceland during the last two years (Grades 9 and 10) of compulsory schooling.

What is school engagement?

Most educational researchers view school engagement[1] as a multidimensional construct (Fredricks, Blumenfeld, & Paris, 2004), which has been defined as "[a] student's active participation in academic and co-curricular or school-related activities, and commitment to educational goals and learning…. It is a multidimensional construct that consists of behavioral (including academic), cognitive, and affective subtypes" (Christenson et. al., 2012, pp. 816-817). As such, school engagement is manifested in active student participation, not only in behavioral terms, such as in participation in academic work (behavioral engagement), but also in emotional and cognitive terms (Li & Lerner, 2011), such as in students' sense of connectedness with their school (emotional engagement), and an interest in learning school material (cognitive engagement). Each subtype alone has been considered a necessary but not a sufficient indicator of school engagement (Li & Lerner, 2013). In other words, school engagement means holistic and integrated active participation (see Dewey, 1913), as compared to mindlessly attending class, or just having good intentions without actively participating in school-related activities.

The importance of intentional self-regulation for adaptive development

Intentional self-regulation (ISR) has been defined as goal-directed behaviors aimed at harmonizing demands and resources in the environment with personal goals (Gestsdottir & Lerner, 2008). ISR allows people to set, prioritize, and obtain long-term goals and promote self-development (Freund & Baltes, 2002). An example of ISR is to intentionally keep a diary to monitor one's progress towards a specific goal and to adjust behavior, and/or the goal, by reflecting on past entries.

---

[1] In the research literature, authors use the terms school engagement and student engagement interchangeably (Libbey, 2004); thus, we make no distinction between these two terms in this study.

The most comprehensive work on the role of ISR in adolescence has been based on the SOC model developed by Baltes and colleagues (1997). According to the SOC model, successful development depends on the effective coordination of selection, optimization, and compensation. The coordination is considered an integrated process of adaptive skills commonly referred to as SOC (Marsiske, Lang, Baltes, & Baltes, 1995). SOC is reflected in a variety of learning situations. For example, when finishing compulsory school in Iceland, admission to the most competitive upper secondary schools is based on final grades in Icelandic language skills, Math, and English. An indication of a student's use of selection, in this situation, would be to identify a need for improvement in English to reach the desired final grade. An example of the same student's use of optimization would be to devote more time to studying English than other subjects where improvement is not perceived to be as important. Finally, if the student's English grades are not improving sufficiently, compensation might occur by getting extra help from a fluent English-speaking friend.

There are multiple examples of specific self-regulatory behaviors that are relevant to school success, including: learning strategies (Zimmerman, 2002), strategies to maintain motivation in education (Wolters, 2003), adaptive help-seeking (Newman, 2002), or practising mnemonic techniques to become a memory expert (Baltes & Baltes, 1990). However, the SOC model captures self-regulatory behaviors that are important to obtain long-term goals across all domains of functioning (Gestsdottir & Lerner, 2008) and the importance of SOC behaviors has been established in several large-scale studies, using samples with adolescents of different ages and cultures, and in relation to various domains of functioning. These studies have contributed to a reliable measurement of SOC among adolescents (Geldhof et al., 2015; Gestsdottir et al., 2015; Gestsdottir, Lewin-Bizan, von Eye, Lerner, & Lerner, 2009), and have demonstrated the importance of SOC skills for promoting positive outcomes (e.g. achievement, confidence, and character) and deterring negative ones (e.g., depression,

substance use; Gestsdottir & Lerner, 2007; Napolitano, Bowers, Gestsdottir, & Chase, 2011; Zimmerman, Phelps, & Lerner, 2008).

The bidirectional relation between school engagement and ISR

The recently developed model of motivational dynamics by Skinner and Pitzer (2012) provides a helpful framework to understand the interaction between school engagement and ISR in relation to school success. The model describes how the context, as well as individual level processes and behaviors, interact to predict learning and achievement (see Figure 1; see also Spencer, Swanson, & Harpalani, 2015). More specifically, learning results from a series of reciprocal feedback effects involving the student's context, the student's self-system processes, and the student's actions.

The reciprocal relation between school engagement and ISR, described in Figure 1 by the concepts engagement and adaptive coping, are at the center of the model and the closest antecedent of learning and achievement.

_____

Insert Figure 1 here

_____

According to this model, both the context and student's own actions can support or undermine engagement and ISR and, as such, academic success. For example, as students move through different school environments that have different constellations of influences (e.g. teachers, peers, courses) some may, at some point, affect their school engagement negatively (Eccles et al., 1993; Simmons, Burgeson, Carlton-Ford, & Blyth, 1987). If so, it is important that the student has the means and flexibility (i.e., ISR) to maintain his or her school engagement from being lowered. Accordingly, the model suggests that a better understanding of the interaction between school engagement and ISR is crucial to identify how to best support young people's learning and academic success.

Prior studies on the relation between school engagment and ISR

There is not a great deal of empirical evidence available that confirms a reciprocal relation between the constructs of school engagement and ISR and none, that we know of, that address this issue from the perspective of intraindividual change. However, there have been two recent variable centered studies that have attempted to assess the proposed reciprocal relation between school engagement and ISR.

In the first study, school engagement (in the form of value, cost, and self-efficacy) and ISR (in the form of rehearsal, organization, elaboration, metacognition, time and study management, and help seeking) was measured among ninth grade students at the beginning and end of a single academic term in the U.S. (Berger & Karabenick, 2011). A cross-lagged structural equation model revealed that school engagement predicted growth in the use of ISR. No support, however, was found for the hypothesis that ISR predicted school engagement. The authors suggested that the temporal interval used in their study (12 weeks) may have accounted for the lack of an effect of ISR skills on changes in school engagement as the use of ISR may require more time before affecting school engagement (Berger & Karabenick, 2011).

In the second study, individual subtypes of school engagement and ISR skills were measured among undergraduate business students in Hong Kong (Ning & Downing, 2010). The measurements were 15 months apart. School engagement was measured by a latent construct labeled affective strategies, and ISR was measured by three latent constructs labeled effort related strategies, comprehension monitoring strategies, and test strategies. A cross-lagged structural equation model revealed that the use of all the ISR skills predicted growth in school engagement to a moderate degree, even when the effects of prior academic achievement were controlled. Contrary to the research hypothesis, and to the findings of Berger and Karabenick (2011), school engagement did not predict growth in ISR skills. The

authors suggested that environmental factors, such as the learning context, may play an undefined role in the relation between ISR and school engagement and called for further research incorporating possible confounding factors into the model (Ning & Downing, 2010).

In sum, whereas the two studies described above separately demonstrated predictive effects in opposite directions between ISR and school engagement, they did not confirm the hypothesized reciprocal effects, but suggested that the findings may have been due to specific circumstances related to the time periods under investigation.

The present study

Despite the importance placed on the reciprocal relation between school engagement and ISR for learning and academic success (Karabenic & Zusho, 2015; Skinner & Pitzer, 2012), such a relation has received minimal empirical support. The current study is a four-wave longitudinal investigation conducted from the beginning of Grade 9 and until the end of Grade 10 (the two last grades of compulsory education) in Iceland. We hypothesized that positive reciprocal effects existed between school engagement and ISR during this time. We tested our hypothesis using a cross-lagged panel model.

The academic context in Iceland makes the last years of compulsory school an especially important focus of study, as it is a time of increased importance for students' future academic prospects. Compared to earlier periods, tenth grade in Iceland is characterized by an increasing urgency to achieve good grades (Sigthorsson, 2008); students with high grades can expect to be admitted to the most competitive upper secondary schools, whereas students with low grades have fewer schools to choose from and run a higher risk of dropping out of school (Blondal, Jonasson & Tannhäuser, 2011).

Finally, one should note that the standard approach to longitudinal analyses has been to analyze data at the between subjects level and the implicit assumption has been that the results are applicable at the within subject-level. This assumption, known as the assumption of

ergodicity, does not hold for some psychological processes (Molenaar, 2004). However, the fact that developmental processes are non-ergodic does not mean that between- subjects analysis is without meri; the extent to which group-level observations reflect person-level phenomena remains a still largely unexplored empirical issue (McClelland, Geldhof, Cameron, & Wanless, 2015). Nevertheless, the current research is limited to between subjects analysis and, as such, the results can only apply reliably to the group-level.

Method

Participants

We randomly selected 20 of the 54 medium- to large-sized schools (>20 Grade 9 students) located in the Reykjavík capital area and the adjacent Reykjanes peninsula. The population in Reykjavík and the Reykjanes area included 66.4% of all Icelandic children. Fifteen out of the 20 selected schools agreed to participate. Each participating school received a book as a gift for their school library for their participation. Two classrooms in each school were selected at random in schools that had more than two classrooms. These 30 classrooms had a total of 625 students. The nested data structure was taken into account when calculating standard errors using a sandwich estimator. A total of 561 parents (90%) gave written consent for their child's participation and 539 (96%) students with parental consent participated at Wave 1 (mean age 14.3, $SD = 0.3$, 46% girls).

Procedure

Participants completed a paper-and-pencil survey during a 40-minute school visit by trained research staff. Standardized instructions were used to increase the probability that data collection was uniformly administered. Students who were absent during the school visit were contacted by e-mail, mail, or phone, and asked to complete and return the survey by mail. During the four waves of measurement, 87%, 80%, 76%, and 84% (respectively) of the 561 participants had complete data on all the school engagement and ISR items. After the last

wave, 61% of the participants had complete data across the four waves of measurement.

Missing data was handled using the full-information maximum likelihood estimation method

(FIML).

Measures

For each measure in the study, a model-based reliability estimate, coefficient ω

(Mcdonald, 1999), was calculated to indicate the proportion of scale variance that was due to

all common factors (Zinbarg, Revelle, Yovel, & Li, 2005). Coefficient ω is analogous to

coefficient α (Reise, 2012); therefore reliability estimates above the 0.70 level were

interpreted as indicators of adequate reliability (Kline, 2011).

School engagement. To measure school engagement, we used the Behavioral-

Emotional-Cognitive School Engagement Scale (BEC-SES) developed by Li and Lerner

(2011, 2013). The BEC-SES consists of three subscales: behavioral, emotional, and cognitive.

Each subscale includes five items on a four-point Likert scale. For each behavioral item,

respondents were asked to rate how often they engaged in specific behaviors using a scale

from 1 (*never*) to 4 (*almost always*)*.* An example of a behavioral engagement item is "How

often do you come to class unprepared (homework unfinished, forget to bring books or other

materials, etc.)?" For the emotional and cognitive items, the respondents were asked to

indicate their agreement on a scale from 1 (*strongly disagree*) to 4 (*strongly agree*). An

example of an emotional school engagement item is "I care about the school I go to?" An

example of a cognitive engagement item is "I think it is important to make good grades" .

A confirmatory bifactor analysis of the BEC-SES has shown evidence of strong

measurement equivalence for the four time points used in the current study. The bifactor

analysis revealed that the covariance among the 15 items was primarily accounted for by a

well-defined single general factor that reflected the common variance among all the items

(Stefansson, Gestsdottir, Geldhof, Skulason, & Lerner, 2015). Consequently, the general

school engagement factor was used in the current study. Coefficient ω for the general BEC-SES in the current study was 0.74, 0.79, 0.74, and 0.75 respectively by wave.

Intentional self-regulation. We operationalized intentional self-regulation (ISR) using a nine item version of the SOC questionnaire (Freund & Baltes, 2002; Gestsdottir & Lerner, 2007). The respondents were asked how they decide what is important for them in life and how they achieve their goals in life. Respondents indicated how closely a list of statements corresponded to their behaviors and/or cognitions on a five point scale ranging from 5 (*Just like me*) to 1 (*Not at all like me)*. An example of a SOC item is "I make every effort to achieve a given goal.". The nine-item version has shown good reliability and validity among Grade 9 students in Iceland (Stefansson, Gestsdottir, & Skulason; 2014). Coefficient ω for the SOC in the current study was 0.70, 0.77, 0.77, and 0.82 respectively by wave.

Covariates. Information about the following covariates was collected at Wave 1: gender, socioeconomic status, mothers' education, fathers' education, mothers' occupation, and fathers' occupation. These variables have been used in large scale international surveys and have shown good reliability among Icelandic adolescents (OECD, 2012b). In addition, information on self-reported grade point average at Wave 1 was collected using an open-ended format.

Data analysis

To assess the presence of a reciprocal relation between school engagement and ISR across the four times of measurement used in this study, a series of factor analyses and structural equation models were estimated using version 7.3 of the Mplus software package (Muthén & Muthén, 1998-2012). To reduce model complexity, parceling was used by aggregating the items of each of the constructs subdimensions (see e.g., Bowers, Wang, Tirrell, & Lerner, 2016). All measures for ISR and school engagement were treated as continuous variables and estimates of latent factors were scaled using the fixed-factor method

of scaling. All models were fit to the data using the maximum likelihood estimator (see Little, 2013).

First, we specified a longitudinal null model that included no expectations of change in the variances or the means of the constructs over time. The longitudinal null model provided an estimate of the overall amount of information contained in the observed data matrix. Second, we allowed the residual variances among the corresponding indicators to be associated over time, and established configural, weak and strong factorial invariance (respectively) using a change of CFI of no more than .01 as a guideline of the assumption that factorial invariance was reasonable (Cheung & Rensvold, 2002).

Next, we fitted a series of structural equation models with varying constraints to the data. In order to examine the theorized reciprocal effects, we built a longitudinal cross-lagged panel model. We built the panel model by starting with freely estimating a minimal set of paths (the bivariate simplex process) and added paths by evaluating information obtained from both modification indices and theory. For each step we conducted a likelihood-ratio test (LRT) to compare the goodness of fit of the competing models. To avoid making a Type 1 error due to the high power of the longitudinal SEM model, we choose a *p*-value less than .001 to determine a significant difference between competing models (see Little, 2013).

Model fit was estimated by evaluating several fit indices: the chi-square statistic, the root mean square error of approximation (RMSEA), and the comparative fit index (CFI). Smaller chi-square and RMSEA values (RMSEA $\leq$ .06), and higher CFI values ($\geq$ .95) indicated a good model fit (West, Taylor, & Wu, 2013). To minimize the risk of making a Type 1 error when using nested data, we ran the final SEM models twice, producing correct standard errors using a sandwich estimator, first based on the class level variation, and again, based on the school level variation. The SEM models showed no appreciable bias in standard errors under either condition.

After the last wave, 61% of the participants had complete data across the four waves

of measurement. No differences in mean scores were found for the ISR variables when

comparing participants with complete data to those with missing data at one or more time

points. However, school engagement at Wave 2 was significantly higher for participants with

complete data compared to those with missing data at one or more time points $t(446) =$

2.97, $p < .001$. Further inspection revealed significant correlations between self-reported

grades, mother's education, father's education, and mother's occupation and missing school

engagement cases at later waves. However, when self-reported grades, mother's education,

father's education, and mother's occupation were included as covariates in the final bivariate

cross-lagged panel model, there were no differences in the study results, indicating limited

bias due to missing data. Consequently, we considered the missing data to be missing at

random (MAR) and used the full-information maximum likelihood estimation method (FIML)

to make use of all the available data in the analysis.

Results

To test the hypothesis that a positive reciprocal relation existed between school

engagement and ISR we built a cross-lagged panel model. As a preliminary caution, we tested

the assumption that our measurements were invariant over time (see Table 1). In addition, we

informed the model building process by reporting the latent model estimates (see Tables 2 and

3). Finally, we included a set of covariates of interest and reported results for the final cross-

lagged panel model (see Figure 2).

Longitudinal invariance of school engagement and ISR

In order to ensure that the same latent construct was being measured at each

measurement occasion, the maximum likelihood estimation method was used to fit a series of

bivariate longitudinal measurement models with varying constraints to the data. The residual

variances of the corresponding indicators were allowed to correlate over time and estimates of

latent factors were scaled using the fixed-factor method of scaling. The configural model gave

an excellent fit to the data, with a very good CFI (.99) and a very good RMSEA (.026). Next,

we constrained each factor loading to be equal across time, these constraints led to a minor

decrease in model fit (ΔCFI = .003), suggesting weak factorial invariance. In addition, we

constrained each intercept to be equal across time. Again, these constraints led to a minor

decrease in model fit (ΔCFI = .005), suggesting strong factorial variance for the bivariate

longitudinal measurement models (see Table 1).

_____

Insert Table 1 here

_____

Descriptive estimates of the latent variables

Latent variable estimates derived from the well-fitting strong invariance model

revealed small mean differences in school engagement (d_{max} = 0.14) and very small

differences in ISR (d_{max} = 0.05) across the four waves of measurement (see Table 2). An

omnibus test of the latent means indicated marginally significant differences ($\Delta\chi2(5)$ = 19.20,

p = 0.002) of school engagement and ISR across the four waves measured. In general, school

engagement tended to be slightly lower during the end of the school year compared to the

beginning of the school year (see Table 2).

_____

Insert Table 2 here

_____

Correlational estimates (see Table 3) revealed very strong correlations (r ranged from

.78 to .82) between measures of school engagement at successive time points. Similarly, the

correlations between ISR at successive time points were strong (.63 to .71). Furthermore, the

analysis revealed strong correlations between school engagement and ISR within each

measurement occasion (.56 to .67), as well as strong correlations between school engagement and ISR across consecutive measurement occasions (.50 to .66).

In sum, the latent means and correlations indicated that the measures of school engagement and ISR were assessing the two constructs in a reliable way across time. In addition, the strong within- and between-construct associations indicated the potential of detecting cross-lagged effects by fitting a structural model to the data.

_____

Insert Table 3 here

_____

Longitudinal structural models

After establishing strong factorial invariance, we used the strong invariance measurement model as a baseline model to further test a series of longitudinal structural equation models in a nested hierarchical manner (see Little, 2013). First, a model of the bivariate simplex process was fit to the data. This first model gave a good fit to the data (see Table 4) but the fit was significantly worse than the baseline model ($\Delta\chi 2(18) = 126.48$, $p < 0.001$). The addition of contextual effects, by allowing measures at Wave 1 to predict measures at Wave 3, and by allowing measures at Wave 2 to predict measures at Wave 4 (i.e., school start/end effects; see Little, 2013) improved the model fit but the model still fitted the data worse than the baseline model ($\Delta\chi 2(14) = 58.21$, $p < 0.001$). The addition of partial cross-lagged effects (i.e., ISR predicting school engagement) further improved the model fit and the model was no longer significantly worse fitting than the baseline model ($\Delta\chi 2(11) = 24.73$, $p = 0.01$). Finally, we fitted the full cross-lagged model, which did not significantly reduce the model fit compared to the baseline model ($\Delta\chi 2(8) = 8.62$, $p = .38$), and gave the best fit to the data of all the longitudinal structural models ($\chi 2 (220) = 342.87$; RMSEA $= .04$; CFI $= .98$). In addition, the full cross-lagged model was compared to a constrained model

where the cross-lagged paths were constrained to be equal across time (see e.g., Engels et al., 2016). The constrained cross-lagged model did not fit the data significantly worse than the full cross-lagged model ($\Delta\chi2(4) = 5.61$, $p = .23$) and was retained as the most appropriate model to answer our research question.

Finally, a set of common covariates (i.e., gender, socioeconomic status, parents' education and occupation, and self-reported grades) were included in the final model to avoid relations between the main study variables being confounded. The covariates were treated as time-invariant and included in the model as predictors for Wave 1 constructs only. Covariates with less than marginally significant effects were excluded from the model one-by-one until only covariates with marginally significant effects ($p < .10$) remained (see Little, 2013). The final model gave a good model fit ($\chi2$ (315) = 556.271; RMSEA = .04; CFI = .97). The final covariate effects (excluded from Figure 2) included small effects of gender ($\beta = 0.10$) and mothers' education ($\beta = 0.15$) on school engagement, and a small effect of fathers' education ($\beta = 0.08$) on ISR. Furthermore, self-reported grades showed a strong effect ($\beta = 0.36$) on ISR, and a very strong effect on school engagement ($\beta = 0.57$).

Figure 2 shows the final standardized structural model with cross-lagged paths constrained to be equal over time. The figure shows that ISR and school engagement were not stable. In fact, ISR and school engagement were significantly less stable during Grade 10 then during Grade 9 ($Wald(2) = 41.99$, $p < .001$). Furthermore, ISR was significantly less stable than school engagement in both grades ($Wald(2) = 9.35$, $p = .009$). Significant cross-lagged effects were found for both ISR and school engagement across the three intervals observed in this study, confirming the hypothesized reciprocal effects between ISR and school engagement during adolescence. The standardized cross-lagged effects from school engagement to later ISR were 0.17, 0.21, and 0.17 respectively by interval. Similarly, although slightly weaker, the standardized cross-lagged effects from ISR to later school

engagment was 0.09, 0.12, and 0.11. respectively by interval. A follow-up test confirmed that the two groups of cross-lagged effects were not significantly different from each other (*Wald*(1) = 12.08, *p* = .72). However, as the effect sizes were small, and research has previously not been able to confirm both of the hypothesized cross-lagged effects in the same model, we decided not to constrain the two groups of cross-lagged effects as equal.

———————————————————

Insert Figure 2 here

———————————————————

In sum, the final model, including several covariate variables, revealed significant cross-lagged effects across the three intervals for school engagement and ISR confirming the hypothesized reciprocal relation between the two constructs. Furthermore, the model-building process revealed a decrease in the stability of both school engagement and ISR from Grade 9 to Grade 10.

Discussion

The promotion of school engagement is widely considered a central focus in the design of effective learning environments and school reform (Shernoff, 2013). This focus stems from growing evidence that links school engagement negatively with school dropout and positively with indicators of academic achievement and well-being (Christenson et. al., 2012). Furthermore, ISR skills are thought to be a key element in the promotion of school engagement (Skinner & Pitzer, 2012). However, little information exists about the hypothesized reciprocal relations between school engagement and ISR skills.

The positive reciprocal relation between school engagement and ISR

The principal aim of this study was to test the hypothesis that a positive reciprocal relation existed between school engagement and ISR during the last two years of compulsory school in Iceland. This hypothesis was confirmed by modeling a cross-lagged relation

between school engagement and ISR during three time intervals that spanned 18 months. The research findings are consistent with the model of motivational dynamics that show school engagement and ISR as mutually reinforcing (see Figure 1; Skinner & Pitzer, 2012). Furthermore, the results revealed a decrease in the stability of both school engagement and ISR during the last two years of compulsory education in Iceland, as well as strong effects of prior academic achievement on school engagement and ISR.

The magnitude of the cross-lagged effects was small across the three time intervals under investigation. However, it should be kept in mind that by definition, the cross-lagged effects were above and beyond the strong autoregressive effects observed. In addition, the fact that the estimates were reciprocal revealed the potential of the cumulative influences that school engagement and ISR may have. In the study of Berger and Karabenick (2011) school engagement weakly predicted the growth of ISR ($\beta = .18$) 12 weeks later. This relation is similar to the weak cross-lagged effects ($\beta = 0.17; 0.21; 0.17$) of school engagement predicting ISR during the three six-month intervals in the current study. However, ISR did not predict school engagement in the study of Berger and Karabenick (2011). The authors concluded that it seemed unreasonable to expect school engagement to change over a 12 week period as suggested by the high stability ($\beta = .82$) of school engagement in their study. The authors hypothesized that the use of ISR skills may require more time before affecting motivational beliefs and called for longitudinal studies that go beyond one semester or even beyond one school year to detect the hypothesized reciprocal relations. Consistent with Berger and Karabenick's (2011) hypothesis described above, our results confirmed that ISR predicted school engagement ($\beta = 0.09; 0.12; 0.11$) over the course of three six-month intervals during the last two school years of compulsory school in Iceland.

School engagement and ISR: Malleable constructs at the end of compulsory school?

As previously reviewed, one of the reasons for the heightened interest in the research on school engagement is that school engagement has been considered to be malleable and, as such, open to contextual influences from parents, teachers, as well as students themselves. However, prior research has concluded that school engagement is very stable between and across school years during adolescence (see e.g., Gottfried, Fleming, & Gottfried, 2001; Skinner & Belmont, 1993). According to the model of motivational dynamics, the apparent stability of school engagement may in fact be the re-creation of stability by the feedback loops between engaged and disaffected actions, on the one hand, and their dependents and antecedents, on the other (see Figure 1; Skinner & Pitzer, 2012). The model-building process in the current study suggested differences in the stability of school engagement across intervals, as the stability during Grade 9 was high and significantly higher than the stability of school engagement during Grade 10. Although these findings need to be confirmed using person-centered analyses, a higher stability in Grade 9 than in Grade 10, could be due to changes in the feedback loops that contribute to the stability, i.e., the multiple negative and positive feedback loops that hold school engagement in place may be more prominent in Grade 9 than Grade 10.

It should be noted that the addition of several education related covariates (i.e., gender, socio-economic status, parents' education and occupation, and self-reported grades) to the cross-lagged panel model highlighted the strong effect of prior academic achievement on later school engagement ($\beta = 0.57$). The strong relation between academic achievement and school engagement, together with the high stability of school engagement in Grades 9 and 10, gives reason to conduct further studies on the relation between academic achievement and school engagement during, and prior to, the period observed in the current study.

Finally, the cross-lagged panel model also revealed differences in the stability of ISR across school years, where, as with school engagement, the stability of ISR during Grade 9 was high and significantly more stable than ISR during Grade 10. A possible decreased stability of ISR and school engagement during Grade 10 would be consistent with theories that present school engagement and ISR as malleable constructs that are open to contextual influences. As previously described, greater expectations for student self-sufficiency and increased urgency of educational goals during Grade 10 seem to coincide with the decreasing stability of both school engagement and ISR. Further person-centered research is needed to explore how contextual influences may relate to school engagement and ISR during the end of compulsory school in Iceland.

Strengths and limitations of the current study and implications for future research

The current study has several limitations and strengths that should be considered when interpreting the results. A key limitation was that the only source of information were self-reports, which are subject to social desirability. Furthermore, this form of data collection may bias the results as the observed correlations between the constructs may be due to common method variance rather than actual relations among the observed variables. However, a considerable strength of this study is that the validity of the school engagement and ISR self-report measures used has been well established with adolescents of different cultures, including Iceland (Geldhof et al., 2015; Gestsdottir et al., 2015; Gestsdottir et. al., 2009; Li & Lerner, 2013; Stefansson et al., 2014; Stefansson et al., 2015).

Another limitation of the current study is that we have only looked at variables and not individual trajectories. The extent to which group-level observations reflect person-level phenomena remains a still largely unexplored empirical issue that future research needs to address. More research is needed to examine the possible existence of subgroups of students that do not fit the normative description of the variable-oriented analysis applied in the current

study. It may be that a subgroup with high levels of school engagement and low levels of ISR skills exists in our data, although the normative trend points to a positive reciprocal relation between the two constructs. Further, person-centered analysis on the development of school engagement and ISR is needed to avoid falsely inferring variable-oriented research results to the person-level.

The somewhat inconsistent results of Berger and Karabenick (2011) and Ning and Downing (2010) regarding the relations between school engagement and ISR indicate that some undefined aspects of the specific learning context may be missing from the models used to simulate the proposed cross-lagged effects. These aspects may be related to the specific period being studied, such as the previously mentioned goal urgency during the last year of compulsory school in Iceland. Furthermore, the size of the cross-lagged effects may be an artifact of the length of the interval chosen, and a shorter or a longer interval might have resulted in a different cross-lagged effect size. Accordingly, future research on the relations between school engagement and ISR should test cross-lagged effects at different periods, with intervals of varying length, and by including measures and designs that are sensitive to the specific learning context (e.g., microgenetic designs; see Bernacki, Nokes-Malach, & Aleven, 2015; Shernoff, 2013).

Finally, it is important to emphasize that predictive effects do not necessarily imply causation. A final limitation worth noting is that the study was conducted in a homogeneous cultural context. The results may, therefore, not be fully applicable with adolescents from other cultures, especially from settings that are more diverse. The research results would benefit from cross-cultural and cross-group validation.

Conclusions

One of the reasons many educational stakeholders are interested in school engagement and ISR skills is that these concepts are expected to be within the sphere of their influence, or

in other words, malleable (Shernoff, 2013). Both constructs are important for academic

achievement and to lower dropout rates, and numerous studies have established uni-

directional effects between school engagement and ISR (see Cleary & Zimmerman, 2012;

Wolters & Taylor, 2012). Scholars have called for a better understanding of the possible

overlap between school engagement and ISR (Boekaerts, 2016; Eccles, 2016) and the

hypothesized reciprocal relation between school engagement and ISR (Pintrich, 2003;

Wolters, 2003). The current research has provided empirical findings to help clarify some of

these issues by providing support for school engagement and ISR skills being highly related,

yet distinct concepts, which mutually reinforce each other during adolescence while

controlling for several important covariates (i.e., academic achievement, gender,

socioeconomic status, mothers' education, fathers' education, mothers' occupation, and

fathers' occupation). The results further suggest that school engagement is, at least at the

inter-individual level, a stable construct during adolescence, and that the use of ISR strategies

may need time before showing detectable relations to school engagement over and above

previous levels of school engagement.

In sum, the current study contributes to a growing body of research that provides

empirical support for school engagement and ISR as mutually reinforcing constructs, which,

in turn, may inform further research on how to best support students' school success and

youth's healthy development.

References

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, 45(5), 369-386. doi:10.1002/Pits.20303

Baltes, P. B. (1997). On the incomplete architecture of human ontogeny - Selection, optimization, and compensation as foundation of developmental theory. *American Psychologist, 52*(4), 366-380. doi:10.1037//0003-066X.52.4.366

Baltes, P. B., & Baltes, M. M. (1990). *Successful aging: Perspectives from the behavioral sciences*. NY: Cambridge University Press.

Berger, J.-L., & Karabenick, S. A. (2011). Motivation and students' use of learning strategies: Evidence of unidirectional effects in mathematics classrooms. *Learning and Instruction, 21*(3), 416-428. doi:http://dx.doi.org/10.1016/j.learninstruc.2010.06.002

Blondal, K., Jonasson, J., & Tannhäuser, A.-C. (2011). Dropout in a Small Society: Is the Icelandic Case Somehow Different? In S. Lamb, E. Markussen, R. Teese, J. Polesel, & N. Sandberg (Eds.), *School Dropout and Completion* (pp. 233-251): Dordrecht: Springer.

Boekaerts, M. (2016). Engagement as an inherent aspect of the learning process. *Learning and Instruction, 43*, 76-83. doi:10.1016/j.learninstruc.2016.02.001

Bowers, E. *P*., Wang, J., Tirrell, J. M., & Lerner, R. M. (2016). A cross-lagged model of the development of mentor-mentee relationships and intentional self-regulation in adolescence. *Journal of Community Psychology*, 44(1), 118-138. doi:10.1002/jcop.21746

Bernacki, M. L., Nokes-Malach, T. J., & Aleven, V. (2015). Examining self-efficacy during learning: variability and relations to behavior, performance, and learning. *Metacognition and Learning, 10*(1), 99-117. doi:10.1007/s11409-014-9127-x

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing

measurement invariance. *Structural Equation Modeling*, 9(2), 233-255.

doi:10.1207/S15328007sem0902_5

Christenson, S. L., Reschly, L. A., & Wylie, C. (Eds.). (2012). *Handbook of research on*

student engagement. NY: Springer

Cleary, T. J., & Zimmerman, B. J. (2012). A cyclical self-regulatory account of student

engagement: Theoretical foundations and applications. In S. L. Christenson, L. A.

Reschly & C. Wylie (Eds.), *Handbook of research on student engagement*. NY:

Springer

Coalition for Psychology in Schools and Education. (2006). *Report on the teacher needs*

survey. Washington, DC: American Psychological Association, Center for Psychology

in Schools and Education.

Dewey, J. (1913). *Interest and effort in education.* Boston, MA: Houghton Mifflin Co.

Eccles, J. S. (2016). Engagement: Where to next? *Learning and Instruction, 43*, 71-75.

doi:10.1016/j.learninstruc.2016.02.003

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., &

Maciver, D. (1993). Development during adolescence - The impact of stage-

environment fit on young adolescents experiences in schools and in families.

American Psychologist, 48(2), 90-101. doi:10.1037//0003-066x.48.2.90

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of

the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109.

Freund, A. M., & Baltes, P. B. (2002). Life-management strategies of selection, optimization,

and compensation: Measurement by self-report and construct validity. *Journal of*

Personality and Social Psychology, 82(4), 642-662. doi:10.1037//0022-3514.82.4.642

Gallup Student Poll. (2015). *Engaged today — ready for tomorrow*. Retrieved from

    http://www.gallupstudentpoll.com/

Geldhof, G. J., Gestsdottir, S., Stefansson, K. K., Johnson, S. K., Bowers, E. *P*., & Lerner, R.

    M. (2015). Selection, optimization, and compensation: The structure, reliability, and

    validity of forced-choice versus Likert-type measures in a sample of late adolescents.

    International Journal of Behavioral Development, 39(2), 171-185.

    doi:10.1177/0165025414560447

Gestsdottir, S., & Lerner, R. M. (2007). Intentional self-regulation and positive youth

    development in early adolescence: Findings from the 4-H study of positive youth

    development. *Developmental Psychology, 43*(2), 508-521. doi:10.1037/0012-

    1649.43.2.508

Gestsdottir, S., & Lerner, R. M. (2008). Positive development in adolescence: The

    development and role of intentional self-regulation. *Human Development*, 51(3), 202-

    224. doi:10.1159/000135757

Gestsdottir, S., Lewin-Bizan, S., von Eye, A., Lerner, J. V., & Lerner, R. M. (2009). The

    structure and function of selection, optimization, and compensation in middle

    adolescence: Theoretical and applied implications. *Journal of Applied Developmental

    Psychology, 30*(5), 585-600. doi:10.1016/j.appdev.2009.07.001

Gestsdottir, S., Geldhof, G. J., Paus, T., Freund, A. M., Adalbjarnardottir, S., Lerner, J. V., &

    Lerner, R. M. (2015). Self-regulation among youth in four Western cultures: Is there

    an adolescence-specific structure of the Selection-Optimization-Compensation (SOC)

    model? *International Journal of Behavioral Development*, 39(4), 346-358.

    doi:10.1177/0165025414542712

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology,* 93(1), 3-13. doi:10.1037//0022-0663.93.1.3

Karabenick, S., & Zusho, A. (2015). Examining approaches to research on self-regulated learning: conceptual and methodological considerations. *Metacognition and Learning, 10*(1), 151-163. doi:10.1007/s11409-015-9137-3

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3. ed.). New York, NY: Guilford Press.

Li, Y. B. (2011). School engagement: What it is and why it is important for positive youth development. In R. M. Lerner, J. V. Lerner, & J. B. Benson (Eds.), *Advances in Child Development and Behavior* (Vol. 41, pp. 131-160). San Diego, CA: Academic press.

Li, Y. B., & Lerner, R. M. (2013). Interrelations of behavioral, emotional, and cognitive school engagement in high school students. *Journal of Youth and Adolescence, 42*(1), 20-32. doi:10.1007/s10964-012-9857-5

Libbey, H. *P*. (2004). Measuring student relationships to school: Attachment, bonding, connectedness, and engagement. *Journal of School Health, 74*(7), 274-283. doi:10.1111/j.1746-1561.2004.tb08284.x

Little, T. D. (2013). *Longitudinal structural equation modeling*. New York, NY: Guilford.

Marsiske, M., Lang, F. R., Baltes, *P*. B., & Baltes, M. M. (1995). Selective optimization with compensation: Life-span perspectives on successful human development. In R. A. Dixon & L. Bäckman (Eds.), *Compensating for psychological deficits and declines: Managing losses and promoting gains* (pp. 35-79). Mahwah, NJ: Erlbaum.

Mcdonald. (1999). *Test theory: A unified treatment.* NJ: Lawrence Erlbaum Associates.

McClelland, M. M., Geldhof, J. G., Cameron, C. E., & Wanless, S. B. (2015). Development and self-regulation. In R. M. Lerner, W. F. Overton, & P. C. M. Molenaar (Eds.),

Handbook of Child Psychology and Developmental Science (7th ed., Vol. 1).

Hoboken, NJ: John Wiley & Sons, Inc.

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: bringing the

person back into scientific psychology, this time forever. *Measurement:*

Interdisciplinary Research and Perspectives, 2(4), 201-218.

doi:10.1207/s15366359mea0204_1

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus User´s Guide* (7 ed.). Los Angeles, CA:

Muthén & Muthén.

Napolitano, C. M., Bowers, E. P., Gestsdottir, S., & Chase, P. (2011). The development of

intentional self-regulation in adolesence: Describing, explaining, and optimizing its

link to positive youth development. *Advances in Child Development and Behavior, 41*,

19-38.

Newman, R. S. (2002). What do I need to do to succeed…when I don't understand what I'm

doing!?: Developmental influences on students' adaptive help seeking. In J. Eccles &

A. Wigfield (Eds.), *Development of Achievement Motivation* (pp. 285-306). San

Diego, CA: Academic Press.

Ning, H. K., & Downing, K. (2010). The reciprocal relationship between motivation and self-

regulation: A longitudinal study on academic performance. *Learning and Individual*

Differences, 20(6), 682-686. doi:10.1016/j.lindif.2010.09.010

OECD. (2012a). *Education at a Glance 2012: OECD Indicators*: OECD Publishing.

OECD. (2012b). *PISA 2009 Technical Report*: OECD Publishing.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in

learning and teaching contexts. *Journal of Educational Psychology, 95*(4), 667.

doi:10.1037/0022-0663.95.4.667

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate*

>   *Behavioral Research, 47*(5), 667-696. doi:10.1080/00273171.2012.715555

Sigthorsson, R. (2008). *Mat í þágu náms eða nám í þágu mats: samræmd próf,*

>   *kennsluhugmyndir kennara, kennsla og nám í náttúrufræði og íslensku í fjórum*

>   *íslenskum grunnskólum.* (Ph.D.), Iceland University of Education, Reykjavik.

Shernoff, D. J. (2013). *Optimal learning environments to promote student engagement* (R.

>   Levesque Ed. 1 ed.). New York, NY: Springer.

Simmons, R. G., Burgeson, R., Carltonford, S., & Blyth, D. A. (1987). The impact of

>   cumulative change in early adolescence. *Child Development, 58*(5), 1220-1234.

>   doi:10.1111/j.1467-8624.1987.tb01453.x

Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom - Reciprocal effects of

>   teacher-behavior and student engagement across the school year. *Journal of*

>   *Educational Psychology*, 85(4), 571-581. doi:10.1037/0022-0663.85.4.571

Skinner, E. A., & Pitzer, J. R. (2012). Developmental dynamics of student engagement,

>   coping, and everyday resilience. In S. L. Christenson, A. L. Reschly, & C. Wylie

>   (Eds.), *Handbook of research on student engagement* (pp. 21-44). New York, NY:

>   Springer.

Spencer, M. B., Swanson, D. P., & Harpalani, V. (2015). Development of the Self. In R. M.

>   Lerner & M. E. Lamb (Eds.), *Handbook of Child Psychology and Developmental*

>   *Science* (7th ed., Vol. 3, pp. 1–44). Hoboken, NJ: John Wiley & Sons, Inc.

Stefansson, K. K., Gestsdottir, S., Skulason, S. (2014). Þróun og mat á réttmæti mælitækis á

>   meðvitaðri sjálfsstjórnun ungmenna. *Sálfræðiritið*, 19, 41-55.

Stefansson, K. K., Gestsdottir, S., Geldhof, G. J., Skulason, S., & Lerner, R. M. (2015). A

>   bifactor model of school engagement: Assessing general and specific aspects of

behavioral, emotional and cognitive engagement among adolescents. *International Journal of Behavioral Development*, 1–10. doi:10.1177/0165025415604056

Tuominen-Soini, H., & Salmela-Aro, K. (2014). Schoolwork engagement and burnout among Finnish high school students and young adults: Profiles, progressions, and educational outcomes. *Developmental Psychology*, 50(3), 649-662. doi:10.1037/a0033898

West, S., Taylor, A., & Wu, W. (2013). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209-231). New York, NY: Guilford.

Wolters, C. A. (2003). Regulation of motivation: Evaluating an underemphasized aspect of self-regulated learning. *Educational Psychologist, 38*(4), 189-205. doi:10.1207/S15326985ep3804_1

Wolters, C. A., & Taylor, D. J. (2012). A self-regulated learning perspective on student engagement. In S. L. Christenson, L. A. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 635-651). New York, NY: Springer

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 19-39). San Diego, CA: Academic Press.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*(2), 64-70.

Zimmerman, S. M., Phelps, E., & Lerner, R. M. (2008). Positive and negative developmental trajectories in us adolescents: Where the positive youth development perspective meets the deficit model. *Research in Human Development, 5*(3), 153-165. doi:10.1080/15427600802274001

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's alpha, Revelle's beta, and McDonald's (omega H) : Their relations with each other and two alternative

conceptualizations of reliability. *Psychometrika, 70*(1), 123-133. doi:10.1007/s11336-

003-0974-7

Tables

Table 1

Model fit statistics for the tests of measurement invariance of ISR and school engagement

across four waves

	χ^2	df	*p*	RMSEA	RMSEA 90% CI	CFI	ΔCFI
Null model	7196.109	312	<.001	---	---	---	---
Configural invariance	259.337	188	<.001	.026	.018;.033	.990	---
Weak invariance	288.789	200	<.001	.028	.021;.035	.987	.003
Strong invariance	334.242	212	<.001	.032	.025;.038	.982	.005

Table 2

Standardized latent means for school engagement and ISR measured at the beginning and the end of Grade 9 (Waves 1 and 2) and Grade 10 (Waves 3 and 4)

	Wave 1		Wave 2		Wave 3		Wave 4	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
ISR	0.00	1	0.05	1.18	0.02	1.19	0.03	1.31
School engagement	0.00	1	-0.14	1.18	-0.01	1.06	-0.11	1.13

Table 3

Latent bivariate correlations derived from the baseline model (strong invariance) across

waves

			ISR				School engagement		
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ISR	Wave 1 (1)	1.00							
	Wave 2 (2)	0.70	1.00						
	Wave 3 (3)	0.61	0.71	1.00					
	Wave 4 (4)	0.53	0.66	0.63	1.00				
School engagement	Wave 1 (5)	0.64	0.59	0.51	0.42	1.00			
	Wave 2 (6)	0.55	0.63	0.59	0.43	0.82	1.00		
	Wave 3 (7)	0.54	0.58	0.67	0.50	0.75	0.78	1.00	
	Wave 4 (8)	0.50	0.54	0.66	0.56	0.69	0.75	0.79	1.00

Note: N = 561; All correlations significant at $p < .001$ level.

Table 4

Summary results from the bivariate model building process

	χ^2	df	$\Delta\chi^2$	Δdf	p	RMSEA	CFI
Baseline (strong invariance)	334.242*	212	-	-	-	.032	.982
Bivariate simplex	460.724*	230	126.482	18	<.001	.042	.966
Contextual	392.456*	226	58.214	14	<.001	.036	.976
Partial cross-lagged	358.967*	223	24.725	11	.010	.033	.980
Full unconstrained cross-lagged	342.866*	220	8.624	8	.375	.032	.982
Full constrained cross-lagged	348.474*	224	5.608	4	.230	.032	.982
Final model with covariates	556.271*	315	-	-	-	.037	.966

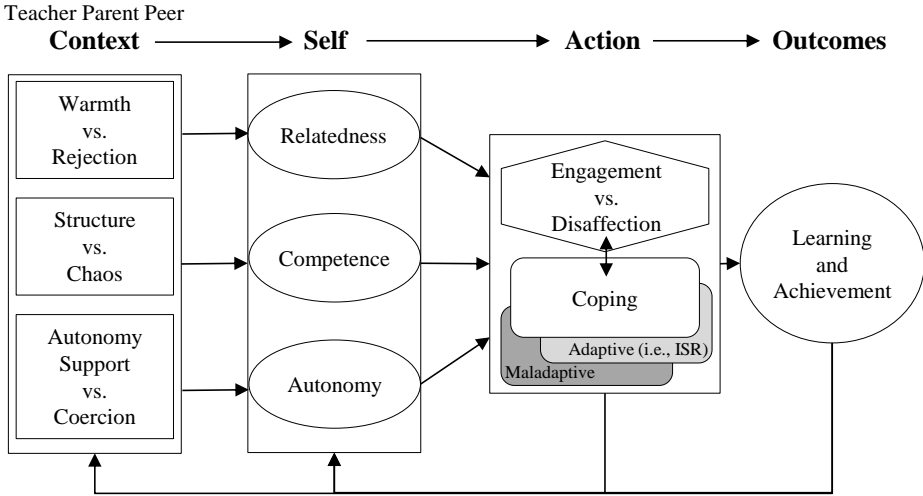*Note:* N = 561; * = significant at $p < .001$ level.

Figure 1

Figure 1. A dynamic model of motivational development organized around student

engagement and disaffection. Figure adapted from Skinner and Pitzer (2012).
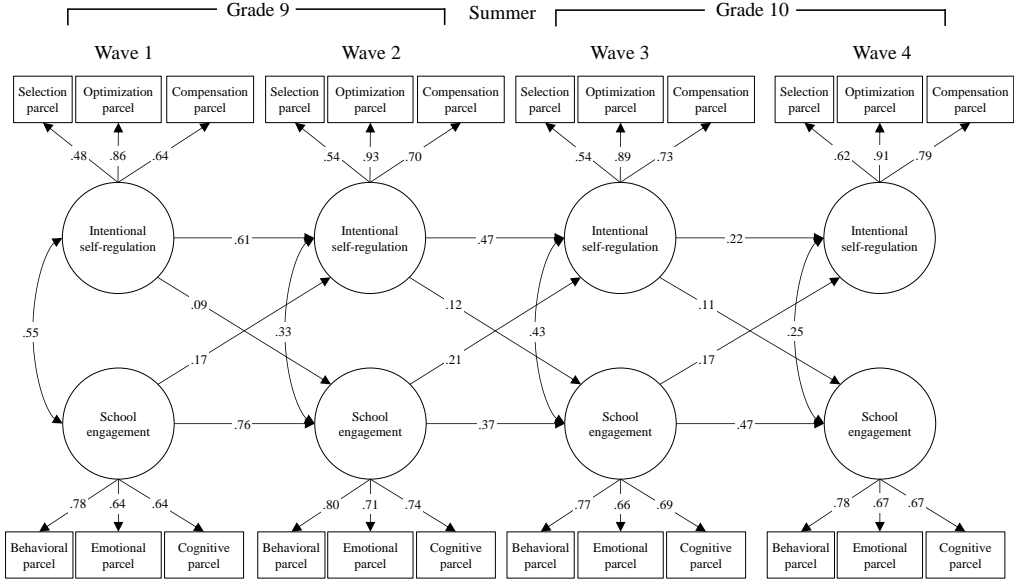
Figure 2

Figure 2. Cross-lagged panel model showing the standardized reciprocal influences between school engagement and ISR at the beginning and end of Grade 9 and Grade 10 in Iceland. The correlating residuals, covariates, and contextual effects are not shown to increase clarity. All the effects in the figure were significant at the *p* < .01 level.