American Psychological Association
Developmental Psychology

0012-1649/19/$12.00
2019, Vol. 55, No. 10, 2090–2101
http://dx.doi.org/10.1037/dev0000764

Complex Problem Solving: Profiles and Developmental Paths Revealed via Latent Transition Analysis

Maida Mustafić
University of Luxembourg

Jing Yu
Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland

Matthias Stadler
University of Luxembourg

Mari-Pauliina Vainikainen
Tampere University and University of Helsinki

Marc H. Bornstein
Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland, and Institute for Fiscal Studies, London, England

Diane L. Putnick
Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland

Samuel Greiff
University of Luxembourg

Complexity is one of the major demands of adolescents' future life as adults. To investigate adolescents' competence development in applying problem-solving strategies in complex environments, we conducted a 2-wave longitudinal study in a sample of Finnish adolescents (11–17 years old; $N = 1{,}959$ at Time 1 and $N = 1{,}690$ at Time 2, 3 years later). In this study, we aimed to: (a) determine the optimal number of strategy use profiles while solving complex problems, (b) determine the number of meaningful developmental paths for each profile, and (c) test the impact of reasoning abilities and learning-related motivational beliefs on the probability that an adolescent with a given strategy use profile will take a given developmental path. Using latent transition analysis, we found 4 meaningful strategy use profiles: *Proficient Explorers*, *Rapid Learners*, *Emerging Explorers*, and *Low-Performing Explorers*. Forty-three percent of the participants were classified as having the same strategy use profile in Time 1 and Time 2. The strategy use of 34% was assessed as having improved between Time 1 and Time 2, while that of 21% was assessed as having declined between Time 1 and Time 2. Verbal reasoning ability and learning-related motivational beliefs predicted whether the developmental path of *Emerging Explorers'* was more likely to remain stable, improve, or decline over time.

Keywords: complex problem solving, strategic behavior, isolated variation strategy, latent transition analysis, reasoning

This article was published Online First July 25, 2019.

Maida Mustafić, Computer-Based Assessment, Institute of Cognitive Science and Assessment, University of Luxembourg; Jing Yu, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland; Matthias Stadler, Computer-Based Assessment, Institute of Cognitive Science and Assessment, University of Luxembourg; Mari-Pauliina Vainikainen, Faculty of Education and Culture, Tampere University, and Centre for Educational Assessment, University of Helsinki; Marc H. Bornstein, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, and Institute for Fiscal Studies, London, United Kingdom; Diane L. Putnick, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health; Samuel Greiff, Computer-Based Assessment, Institute of Cognitive Science and Assessment, University of Luxembourg.

Matthias Stadler is now at the Department of Education and Educational Psychology, University of Munich.

Samuel Greiff is one of two authors of the commercially available COMPRO-test that is based on the multiple complex systems approach and that uses the same assessment principle as MicroDYN. However, for any research and educational purpose, a free version of MicroDYN is available. Samuel Greiff receives royalty fees for COMPRO.

This research was supported by a fellowship and a grant awarded to Samuel Greiff by the Fonds National de la Recherche Luxembourg (ASKI21; TRIOPS); Jing Yu, Diane L. Putnick, and Marc H. Bornstein were supported by the Intramural Research Program of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health; and Marc H. Bornstein by an International Research Fellowship in collaboration with the Centre for the Evaluation of Development Policies (EDePO) at the Institute for Fiscal Studies (IFS), London, United Kingdom, funded by the European Research Council (ERC) under the Horizon 2020 research and innovation programme (Grant agreement 695300-HKADeC-ERC-2015-AdG).

Correspondence concerning this article should be addressed to Maida Mustafić, who is now at the Institute Humans in Complex Systems, School of Applied Psychology, University of Applied Sciences and Arts Northwestern Switzerland, Riggenbachstrasse 16, 4600 Olten, Switzerland. E-mail: maida.mustafic@fhnw.ch

Developmental researchers and educational stakeholders alike recognize the importance of adolescents' mastery of problems in complex environments (complex problem solving; CPS) for meeting the challenges awaiting them now and in the future. Like creative thinking, collaboration, communication skills, and adaptability, CPS is a 21st-century skill that is considered as necessary and relevant for adolescent's successful performance across a range of school subjects and adolescent biographical development in the long term (OECD, 2014). Because of its predictive validity with respect to educational and work achievement (e.g., Danner, Hagemann, Schankin, Hager, & Funke, 2011; Greiff, Stadler, Sonnleitner, Wolff, & Martin, 2015; Kyllonen, Anguiano Carrasco, & Kell, 2017; Schweizer, Wüstenberg, & Greiff, 2013), CPS was incorporated into the Program of International Student Assessment (PISA), the largest educational assessment program worldwide, with approximately half a million participants each cycle (OECD, 2014). Because CPS is relevant for adolescents' biographies and is, because of its generalizability across domains, applicable to studies of children's and adolescents' behavioral development, we have a potential, but also the need to understand and to foster the successful development of CPS.

Adolescents' Complex Problem Solving Strategies

The definition of CPS reflects the challenge of preparing adolescents for a world that is constantly evolving in complexity and dynamic in change. Buchner (according to Frensch & Funke; Buchner, 1995, p. 14) defines CPS as

> the successful interaction with task environments that are dynamic (i.e., change as a function of the user's interventions and/or as a function of time) and in which some, if not all, of the environment's regularities can only be revealed by successful exploration and integration of the information gained in that process.

CPS requires the active creation, selection, and integration of knowledge rather than solving problems by applying previous knowledge or familiar strategies. In contrast to more domain-specific competencies such as reading or science problem-solving, CPS can be considered a rather domain-general cognitive ability that does not require much domain-specific factual knowledge (Greiff & Wüstenberg, 2014). Because only little content knowledge is required for successful complex problem solving, planned multistep interventions and strategic exploration are core aspects of CPS that differentiate CPS from other cognitive abilities, including reasoning—that is, drawing inferences and conclusions from one's own thinking based on epistemic constraints (Raven, 2000; Stadler, Niepel, & Greiff, 2019; Wüstenberg, Greiff, & Funke, 2012).

Strategic behavior, that is, approaching problems with a higher-order plan to achieve a particular goal under dynamic and nontransparent conditions, has been identified as highly relevant for successful problem solving (Chen & Klahr, 1999; Greiff, Wüstenberg, & Avvisati, 2015; Kuhn et al., 1995; Wüstenberg, Stadler, Hautamäki, & Greiff, 2014). One basic, but important behavioral strategy associated with successful CPS is the strategy of isolated variation (e.g., Chen & Klahr, 1999; Tschirgi, 1980). Isolated variation is defined as the method of manipulating one variable in the problem environment while holding all other variables constant (Chen & Klahr, 1999). Mastery of the strategy of isolated variation is a precondition for (a) the development of more sophisticated strategies involving the coordination of multiple variables, and (b) the different phases of scientific thinking: inquiry, analysis, inference, and argument (e.g., Kuhn, 2010; Kuhn et al., 1995). It is involved in hypothesis generation, experimental design, and evidence evaluation and is applicable domain-generally (Zimmerman, 2000).

Regarding CPS performance, Greiff, Wüstenberg, and Avvisati (2015) found that adolescents who used the isolated variation strategy in CPS tasks performed better on specific tasks and had higher overall CPS test scores. These findings stress the importance of taking a strategic approach to CPS as they indicate that adolescents who used the isolated variation strategy most likely understood the basic principles of strategic behavior and were able to transfer their strategic behavior to other tasks. Although the importance of applying a strategic approach in CPS has been shown, research about how adolescents can be differentiated on the basis of their ability to take a strategic approach in CPS is scarce. One recent effort to differentiate between patterns of strategic behavior using a cross-sectional person-centered approach yielded remarkable qualitative differences between adolescents (Greiff, Molnár, Martin, Zimmermann, & Csapó, 2018). In that latent class analysis study, adolescents were classified into six groups that differed significantly with respect to overall CPS performance and duration of task exploration behavior. Those results provided valuable insight into differences in adolescents' use of the isolated variation strategy, but they were cross-sectional and did not contribute to our knowledge about developmental paths to distinct strategic behavior profiles or any factors influencing those paths.

Factors Predicting the Development of Strategic Behavior in CPS

Domain-independent, abstract, and rule-independent strategic behavior can be observed from elementary school on and it develops from childhood to adolescence (e.g., Zimmerman, 2000). Adolescents at the age of 16 already adapt their use of the isolated variation strategy in a task depending on its effectiveness and on their own intelligence. In line with the current understanding of intelligence as the "ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, [and] to overcome obstacles by taking thought" (Board of Scientific Affairs of the American Psychological Association; Neisser et al., 1996, p. 77), the more intelligent adolescents are, the more flexibly they are able to use the strategy (Lotz, Scherer, Greiff, & Sparfeldt, 2017). Researchers investigating longitudinal predictors of CPS, however, have focused more on CPS performance than on strategic behavior during CPS (e.g., Frischkorn, Greiff, & Wüstenberg, 2014; Greiff, Wüstenberg, Goetz, et al., 2015). A longitudinal study on Finnish adolescents ($N = 2{,}021$), for example, revealed that reasoning was a precursor of CPS–unlike working memory, which had only a small impact on CPS (Greiff, Wüstenberg, Goetz, et al., 2015). These findings support the idea that CPS involves reasoning and, to a smaller degree, working memory in the initial stages of development and develops into a higher-order cognitive skill in adolescence. This finding is in line with studies showing that early cognitive information-processing levels predict future complex cognitive performance. For example, in preverbal infants, habituation efficiency is a small but significant predictor, above and beyond temperament and experience, of more complex cognition in childhood (e.g., Bornstein et al., 2006). From childhood until adolescence, cognitive abilities continue to develop in a cascade, in that processing speed and working memory are significant precursors of the reasoning ability (Fry & Hale, 1996; Nettelbeck & Burns, 2010). Given that reasoning predicts CPS performance, and strategic behavior is related to CPS performance, it is likely that reasoning is also predictive of the development of strategic behavior. Therefore, in the present study, we hypothesized that greater reasoning ability would be associated with the development of better strategy use during CPS. We differentiated between three facets of the reasoning ability: (a) scientific reasoning, defined as a set of thinking skills involved in inquiry, experimentation, causal design, concept formation and similar (Dunbar & Klahr, 2012); as well as (b) figural and (c) verbal reasoning, defined as the process of drawing inferences in problem situations that are based on (a) abstract or (b) linguistic principles (e.g., Polk & Newell, 1995).

Besides cognitive abilities, motivational abilities are assumed to significantly predict complex human behavior (e.g., Dörner & Güss, 2013). Ample empirical evidence confirms that motivational beliefs do have an impact on cognitive achievement (e.g., Eccles & Wigfield, 2002), although these effects are small or medium in size (Gagné & St. Père, 2001). In particular, learning-related motivational beliefs, which can be defined as motivational beliefs to guide and support the effective use of cognitive abilities (Hautamäki et al., 2002), were found to influence adolescents' achievement, especially in the face of difficulty or failure. As indicated by several meta-analyses, perceiving success as the result of one's personal control of effort and learning affects successful performance as well as persistence in the face of failure, whereas the goal to surpass one's peers is associated with poor achievement, negative affect, and avoidance of challenges (e.g., Dweck, 1986; Findley & Cooper, 1983; Hulleman, Schrager, Bodmann, & Harackiewicz, 2010; Rawsthorne & Elliot, 1999). The effect of focusing attention on effort and the use of apt strategies was shown to be particularly strong in complex tasks (Utman, 1997). Relating to CPS, Rudolph, Niepel, Greiff, Goldhammer, and Kröner (2017) showed that adolescents' confidence in being able to solve CPS tasks was one factor related to eventual CPS performance. Also, longitudinal studies confirmed that initial levels of learning-related motivational beliefs assessed in primary school predicted CPS performance at the end of the 6th grade (Vainikainen, Wüstenberg, Kupiainen, Hotulainen, & Hautamäki, 2015). As learning-related motivational beliefs have been shown to have an impact on CPS, we reasoned that learning-related motivational beliefs at Time 1 would likely predict the development of adolescents' use of strategies. The effect might be based on the impact of those beliefs on the likelihood that adolescents put effort into learning in general. Therefore, for the present study, we included cognitive and motivational variables to test whether reasoning and learning-related motivational beliefs shape the development of adolescents' strategy use during CPS over time.

The Present Study

Because of its implications for adolescents' learning and development, it is important to investigate the development of adolescents' use of strategies in complex environments. Gaining a better understanding of adolescents' strategic approach over time and identifying factors that predict the development of the approach will make researchers and practitioners more able to foster students' success in these environments. The present two-wave longitudinal study was, therefore, designed to take a person-centered latent mixture modeling approach to reveal the pathways of development of strategic behavior during CPS in adolescents. Thus, the present study aimed to: (a) determine the optimal number of strategy use profiles for the assessment of adolescents, (b) determine the number of meaningful developmental paths for each strategy use profile, and (c) test the impact of reasoning abilities and learning-related motivational beliefs on the probability that an adolescent with a given strategy use profile will take a given developmental path. This will be the first longitudinal, differentiated, and dynamic picture of the development and determinants of differences in strategic behavior during CPS in adolescents.

Method

Participants

The participants were adolescents from a municipality in Finland who took part in a longitudinal educational assessment project ("Learning to learn in basic education schools") conducted regularly and carried out by the Centre of Educational Assessment (CEA) at the University of Helsinki in Finland in 2013 and 2016. All schools in the municipality were recruited. According to an ethical review of our research protocol, we informed the adolescents' parents about the assessment and the adolescents gave their consent on participation by agreeing on taking the test.

All participants were enrolled in "comprehensive schools," which are the mandatory part of the educational system in Finland, namely, from 1st to 9th grade (for children/adolescents aged 7 to 16; Finnish peruskoulu, Swedish grundskola, literal translation: "basic school"). The same CPS test was administered in the 6th grade in 2013 (Time 1; N = 1,959; 49,5% girls) and in the 9th grade 2016 (Time 2; N = 1,690; 52% girls) assessments (see further descriptions of the sample at Time 1 and Time 2 in Table 1). The investigated age range represents a developmental period in which adolescents still develop their CPS skills and are sensitive to training and intervention (Molnár, Greiff, & Csapó, 2013). Full information maximum likelihood estimation (FIML; Collins, Schafer, & Kam, 2001) was used in M*plus* Version 7.4 (Muthén & Muthén, 2010) to handle missing data.

Procedure

The Education Department of the municipality organized the ethical approval of the project "Learning to learn in basic education schools" and the used measures. The National Institutes of Health and Welfare and the Ministry of Education and Culture, both in Finland, approved the research protocol that was used. Before the assessment, the Education Department informed the school principals in detail and provided them with an information

Table 1
Sample Characteristics in 2013 (Time 1) and in 2016 (Time 2)

Descriptive statistics	Time 1		Time 2	
	Girls	Boys	Girls	Boys
N (%)	971 (49.5%)	988 (50.5%)	817 (48.5%)	873 (52,0%)
M_{age}	12.19	12.21	15.24	15.23
SD_{age}	.42	.42	.44	.45
$\text{Range}_{\text{age}}$	11 to 14	11 to 14	14 to 17	15 to 17

letter for the parents. The parents had the opportunity to receive further information on the study and also to withdraw their children's participation, if desired.

At both time points, the assessments took place on school days as a part of regular school activities; the participants did not receive external rewards for participation. Homeroom teachers received detailed written instructions to administer the test sessions. The teachers were also asked to log in to the online assessment system to get familiar with the content before the test sessions. All tests were administered using tablet computers and the members of the assessment team visited some of the sessions for quality monitoring reasons. The participants were assessed via an online test battery consisting of three blocks of 45-min sessions. In the first two 45-min sessions, the participants completed cognitive assessment tasks, learning-related motivational beliefs questionnaires, and a personal background questionnaire. The cognitive tasks comprised reading comprehension, mathematical thinking skills, and reasoning tasks, of which only the reasoning tasks were used in this study. The motivational beliefs questionnaires consisted of several scales from different theoretical origins, of which one questionnaire was used in this study. The order of presentation of tasks and questionnaires was counterbalanced so that tasks of different domains were distributed evenly throughout the session. Self-report questionnaires were divided up into different sets that were presented between cognitive tasks.

In the final 45-min session, the participants watched a short video built in the assessment platform to learn the principles of the CPS tasks. Then, they completed the online CPS tasks and a questionnaire assessing test anxiety.

The data from this panel study were used also for other publications. The general design of the study is described in Vainikainen (2014). The research questions and analyses presented in this manuscript have not been addressed previously and are unique.

Measures and Scoring

Isolated variation strategy. At Time 1 and Time 2, we assessed the participants' use of the isolated variation strategy using tasks based on the MicroDYN framework (Greiff, Wüstenberg, & Funke, 2012). MicroDYN tasks are a reliable and valid computer-based measure of complex problem solving ability; they were used to measure problem-solving performance in the PISA 2012 survey (OECD, 2014). In the MicroDYN CPS tasks, adolescents work on nontransparent and dynamically changing problems by exploring the impact of three input variables on three output variables within fictitious scenarios (e.g., in a "creating a perfume," "feeding a cat," or "driving a moped" scenario; see also Greiff et al., 2012; Schweizer et al., 2013; Wüstenberg et al., 2012). Each MicroDYN task consists of two phases, a knowledge acquisition phase (maximum time allotted: 180 s) and a knowledge application phase (maximum time allotted: 90 s). Six MicroDYN CPS tasks for which isolated variation was the optimal strategy were chosen for this study. The isolated variation strategy was scored only during the knowledge acquisition phase.

For example, in the knowledge acquisition phase of the MicroDYN task called "Perfume" (see Figure 1), the participants were to adjust the levels of three input variables, which represented the fictitious ingredients of a perfume ("Norilan," "Miral," and "Karumin"), to observe changes in the output variables, which represented the resulting scents ("Fresh," "Fruity," and "Flowery"). The participants were to adjust the levels of input variables in any way they deemed necessary to discover the impact of the ingredients on the resulting scents. After the participants had adjusted the level of at least one ingredient or combined up to three ingredients, they were to click on "Apply" to observe the single or combined effect on the resulting scent. Use of the isolated variation strategy was operationalized as manipulation of only one ingredient at a time before clicking Apply. Isolated variation was coded by a computer algorithm that registered how many variables were manipulated at a time before clicking Apply. A participant's score was either 0 (*no use of isolated variation strategy at all*), 1 (*partial use of isolated variation strategy*; *isolated variation strategy was applied to some, but not all input variables*), or 2 (*isolated variation strategy was applied to all input variables*). Thus, the indicators for the latent transition analysis were ordered categorical indicators with three categories (0, 1, and 2). Table 2 depicts the category frequencies for isolated variation strategy use at Time 1 and Time 2.

Scientific reasoning ability. We measured scientific reasoning ability at Time 1 using the "control of variables task" (Hautamäki, 1984), a modified version of Shayer's (1979) science reasoning task called "Pendulum." We presented the participants two scenarios in the realm of Formula 1 races to be compared based on the effect of four variables: driver, car, tires, and track. The participants were asked to judge whether the single effect of the driver, car, tires, or track could be concluded from comparisons of the effects described in tables ("Based on the information given in the table, can you make a conclusion on the effect of the driver on the end result?"). Then, two comparison sets were shown to the participants. The first set was a table with a prefilled first row; the participants had then to decide on the properties of the second row to be able to make the conclusions specified in the task. In the second comparison set, the table was empty and the participants were asked to construct a hypothetical comparison to be able to study the effects specified in the task. In total, there were four decisions on the effects of single variables and two comparison sets to be completed. The items were coded dichotomously, and we calculated a mean score from the six coded items ($N = 1{,}907$; $M = 2.11$, $SD = 1.61$). The reliability of the test at T1 was acceptable ($\alpha = .68$).

Verbal reasoning ability. We measured verbal reasoning ability at Time 1 using five items from the "missing premises task" from the Ross Test of Higher Cognitive Processes (Ross & Ross, 1979). We presented the participants with the first of two premises and a conclusion, and they selected the second premise from five alternatives that would make the conclusion valid. (e.g., "Conclusion:

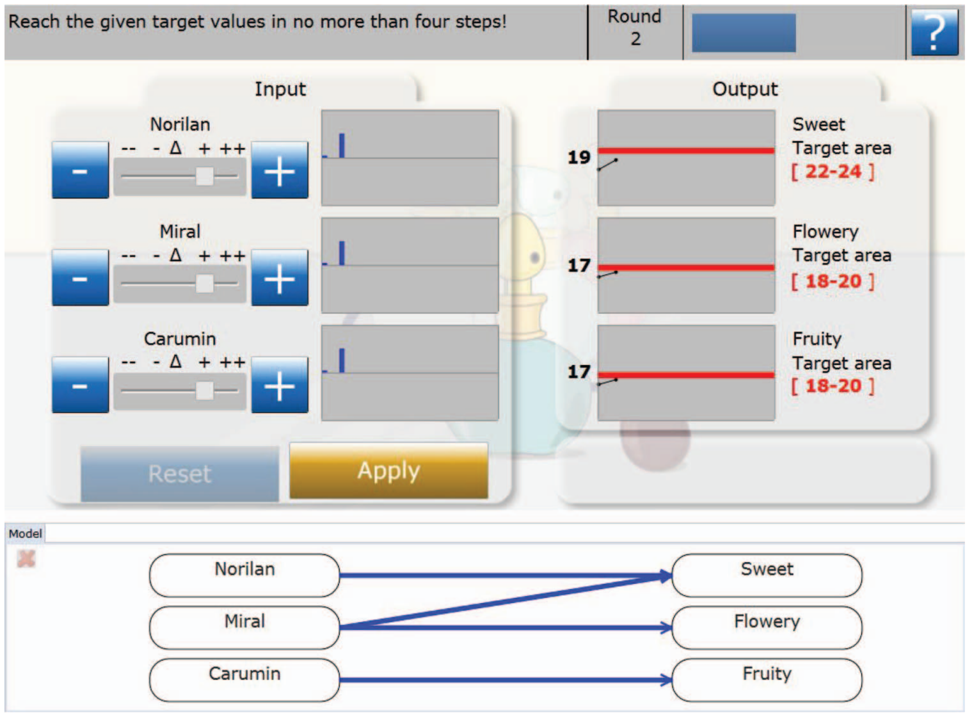

Figure 1. Screenshot of the MicroDYN CPS task "Perfume" during the knowledge application phase. "Norilan," "Miral," and "Carumin" represented fictitious ingredients of a perfume (input variables). "Fresh," "Fruity," and "Flowery" represented the resulting scents (output variables). See the online article for the color version of this figure.

Lake Saimaa is too cold for swimming. First fact: The temperature of Lake Saimaa is 5 centigrades. Which of the following is the second fact if the conclusion is valid? Most lakes are too cold for swimming.; It is wintertime.; Five degree water is too cold for swimming.; Lake Saimaa is always cold.; Swimming in cold water is no fun."). We scored the participants' responses to each item dichotomously as correct or incorrect and averaged them to produce a mean score (N = 1,936; M = 2.03, SD = 1.20). Because of the small number of items, the reliability of the scale at T1 was rather low (α = .46).

Figural reasoning ability. A subsample of participants completed a paper-based figural reasoning task at Time 1. Figural reasoning was measured using the classical Piagetian water-level task for understanding horizontal and vertical axes (Hautamäki, 1984; Piaget & Inhelder, 1956). We presented a picture of eight empty bottles to the adolescents. One of the bottles was standing upright in a vertical position, and the rest were inclined by 45, 90, 135, 320, 270, 225, and 180° from a vertical position, respectively. The task was to draw a line indicating the water level inside the bottle and color the area filled with water when each bottle was half full. Each bottle was scored dichotomously as correct or incorrect, and then a mean score was calculated for the eight coded items (N = 402; M = 5.26, SD = 2.28). The reliability of the measure was good at T1 (α = .83).

Learning-related motivational beliefs. We assessed the participants' learning-related motivational beliefs using the "control

Table 2
Absolute and Relative Category Frequencies for the Six CPS Tasks for Time 1 and Time 2

	Time 1				Time 2			
Categories	n	0	1	2	n	0	1	2
CPS task								
Task 1	1,704	557 (.33)	573 (.34)	574 (.34)	1,377	471 (.34)	320 (.23)	586 (.43)
Task 2	1,705	591 (.35)	539 (.32)	575 (.34)	1,370	496 (.36)	265 (.19)	609 (.44)
Task 3	1,704	623 (.37)	495 (.29)	586 (.34)	1,364	489 (.36)	277 (.20)	598 (.44)
Task 4	1,702	777 (.46)	359 (.21)	566 (.33)	1,351	538 (.40)	163 (.12)	650 (.48)
Task 5	1,700	791 (.47)	322 (.19)	587 (.34)	1,344	566 (.42)	153 (.11)	625 (.46)
Task 6	1,704	785 (.46)	311 (.18)	608 (.36)	1,313	565 (.43)	134 (.10)	614 (.47)

Note. CPS = complex problem solving; 0 = no use of isolated variation strategy; 1 = partial use of isolated variation strategy; 2 = consistent use of isolated variation strategy.

expectancy scale" from the Finnish learning-to-learn instrument (see also Vainikainen et al., 2015). The instrument is based on the Finnish learning to learn model (LTL model; e.g., Hautamäki et al., 2002), which was derived from attribution and social–cognitive theory (Bandura, 2006; Rotter, 1990). The three items of the scale assess whether adolescents perceive themselves as capable of learning and whether they perceive learning as something controllable and changeable. The items ("If I don't master yet what I would like to, I know I will be able to learn it," "If I don't yet master something I need to, I will learn it easily enough," "If learning something is important to me, I know I'll learn it") were rated on a 7-point Likert scale ranging from 1 (*not true at all*) to 7 (*very true*). We computed a mean score of the three items at T1 (N = 1,940; M = 5.30, SD = 1.01; α = .61).

Analysis Plan

We conducted a latent transition analysis (LTA) in M*plus* Version 7.4 (Muthén & Muthén, 2010); this type of analysis is a longitudinal extension of latent class analysis (LCA), and both are specific cases of latent mixture modeling. Latent mixture modeling is a person-centered approach to modeling heterogeneity by classifying individuals into subgroups with similar characteristics. LCA is a multivariate statistical model, according to which a latent variable can be inferred from a set of categorical indicators in the way that individuals in each class show common behavior patterns, that is, the individuals are members of a particular latent class. LTA (also known as hidden Markov modeling) is a longitudinal mixture model in which individuals are allowed to transition between classes with a particular probability. Lanza, Patrick, and Maggs (2010) suggest to name the patterns latent "classes" in LCA, "profiles" in latent profile analysis (LPA), and latent "statuses" in LTA. We kept the term "profile" in this LTA instead of "status" because we found it the more specific and descriptive term to describe different CPS strategy use patterns.

In LTA, latent profile membership probabilities, transition probabilities, and item response probabilities are estimated. Latent profile membership probability is the probability that a participant will be classified as having a given profile at a certain time point. Transition probability is the probability that a participant will change profile membership over time, and item response probability is the conditional probability that a participant will always use the isolated variation strategy, under the condition that one has been classified as having a particular profile at Time 1. The mathematical model for LTA is presented in Appendix (see Collins & Lanza, 2010, for a more comprehensive introduction to LTA).

In the present study, we examined multivariate behavior profiles and transitions from Time 1 (2013) to Time 2 (2016). Based on statistical criteria, parsimony, and interpretability, we selected the optimal number of profiles. For the sake of interpretability, the number of latent profiles per time point is ideally as parsimonious as possible and set to be equal over time. Statistical models should estimate no more parameters than absolutely necessary to represent the data adequately (Box & Jenkins, 1970). As for statistical criteria, smaller values of relative fit indices (Akaike's Information Criterion, AIC, Akaike, 1987; and Bayesian Information Criterion, BIC, Schwarz, 1978) are preferred. Fit indices indicate whether the distribution of the data corresponds to the population model and represents a more optimal balance between model fit and parsimony. However, indices like AIC and BIC are more useful for ruling out alternative models than showing which models are absolutely valid. To ensure that profiles are equally interpretable across time points, we held item response probabilities equal across time. This is a common procedure in the context of LTA to enhance the parsimony and interpretability of the data (see also Rinne, Ye, & Jordan, 2017). We then compared the models with constrained response probabilities (measurement invariant models) to models with freely estimated item response probabilities to evaluate the fit of the more parsimonious models. Also, because mixture models are susceptible to local solutions, multiple runs with different sets of random values were performed to ensue rather a global than a local fit is obtained (see also Feldman, Masyn, & Conger, 2009; Hipp & Bauer, 2006). In all the runs with different sets of starting values the solutions were replicated.

Finally, we included reasoning and learning-related motivational beliefs at Time 1 as covariates in the model. To test the effects of reasoning and learning-related motivational beliefs on the latent transitions over time multinomial logistic regression was used. The basic idea was that the interaction of the covariates with the item response probabilities at Time 1 would predict the item response probabilities at Time 2. In that, each covariate has k-1 (k is the number of profiles) different complementary effects when one profile is compared against another profile. Odds ratios (*OR*s) reflect the change in likelihood of membership in the target profile when compared with the reference profile for each unit increase in the predictor. An *OR* of 2 indicates that for each unit change in the predictor, the participants are two times as likely to be classified as belonging to the target profile in comparison to the reference profile; an *OR* smaller than 1 represents a negative logistic regression coefficient and indicates that the likelihood of belonging to the target decreased. For example, an *OR* of .5 means the likelihood of membership to the target profile compared with the comparison profile is lessened by 50% per unit increase in the predictor (see also Kam, Morin, Meyer, & Topolnytsky, 2016). We included the covariates separately into the LTA models to investigate the separate effects of scientific, verbal, and figural reasoning as well as learning-related motivational beliefs (see also Collins & Lanza, 2010).

Results

Goal 1: Determine the Optimal Number of Strategy Use Profiles

We assessed model identification for seven measurement invariant and variant LTA models. Table 3 shows the model fit information that was used to determine the final model. For reasons of interpretability and parsimony, we preferred the measurement invariant (constrained) models, although the freely estimated (unconstrained) models showed a slightly better model fit. The benefit of the measurement invariant models is that the nature and meaning of the latent profiles is constant across time and, therefore, can be interpreted. Of the measurement invariant models, the four-class model was superior to the two-class and three-class models with respect to AIC and BIC. Moreover, the four-class model was superior to the five-, six-, seven-, and eight-class models with

Table 3
Fit Indices of Unconstrained and Constrained LTA Models Differing in the Specified Number of Latent Profiles

Index	Two	Three	Four	Five	Six	Seven	Eight
			Unconstrained LTA models				
Log likelihood	−14079.53	−13313.71	−12988.30	−12767.75	−12572.83	−12399.25	−12255.73
Free parameters	99	224	399	624	899	1244	1599
AIC	28357.06	27075.42	26774.60	26783.50	26943.65	27264.51	27709.46
BIC	28911.80	28330.58	29010.36	30280.02	31981.11	34105.10	36669.30
Entropy	.81	.75	.75	.78	.78	.80	.80
			Constrained LTA models				
Log likelihood	−14547.57	−13632.82	−13328.44	−13213.17	−13112.16	−12979.99	−12901.95
Free parameters	27	44	63	84	107	216	351
AIC	29149.14	27353.64	26782.89	26594.34	26438.32	26391.97	26505.89
BIC	29300.43	27600.19	27135.90	27065.02	27037.89	27602.31	28472.68
Entropy	.75	.70	.70	.66	.68	.70	.71

Note. LTA = latent transition analysis; AIC = Akaike information criterion; BIC = Bayesian information criterion.

respect to entropy and parsimony. Thus, we decided to keep the four-class model for further interpretation.

Figure 2 shows the item response probabilities of the profiles and graphically depicts the four different profiles. As the item response probabilities were constrained to be equal across the two time points, the eventual profiles are identical for Time 1 and Time 2. Participants classified as *Proficient Explorers* used the isolated variation strategy across the six CPS tasks consistently (>80%). Participants classified as *Rapid Learners* were not likely (<50%) to apply the isolated variation strategy in the first three CPS tasks, but more likely (>50%) to apply the strategy in the last three CPS tasks. Participants classified as *Emerging Explorers* applied the isolated variation strategy in the first few CPS tasks, but then reverted to only partially using the isolated variation strategy as the predominant strategy in the later, more difficult tasks. Finally, participants classified as *Low-Performing Explorers* were highly likely (>60%) to fail to use the isolated variation strategy to solve the six CPS tasks.

Goal 2: Determine the Number of Meaningful Developmental Paths and the Transition Probabilities

Our study with two time points and four profiles of isolated variation strategy use (*Proficient Explorers*, *Rapid Learners*,

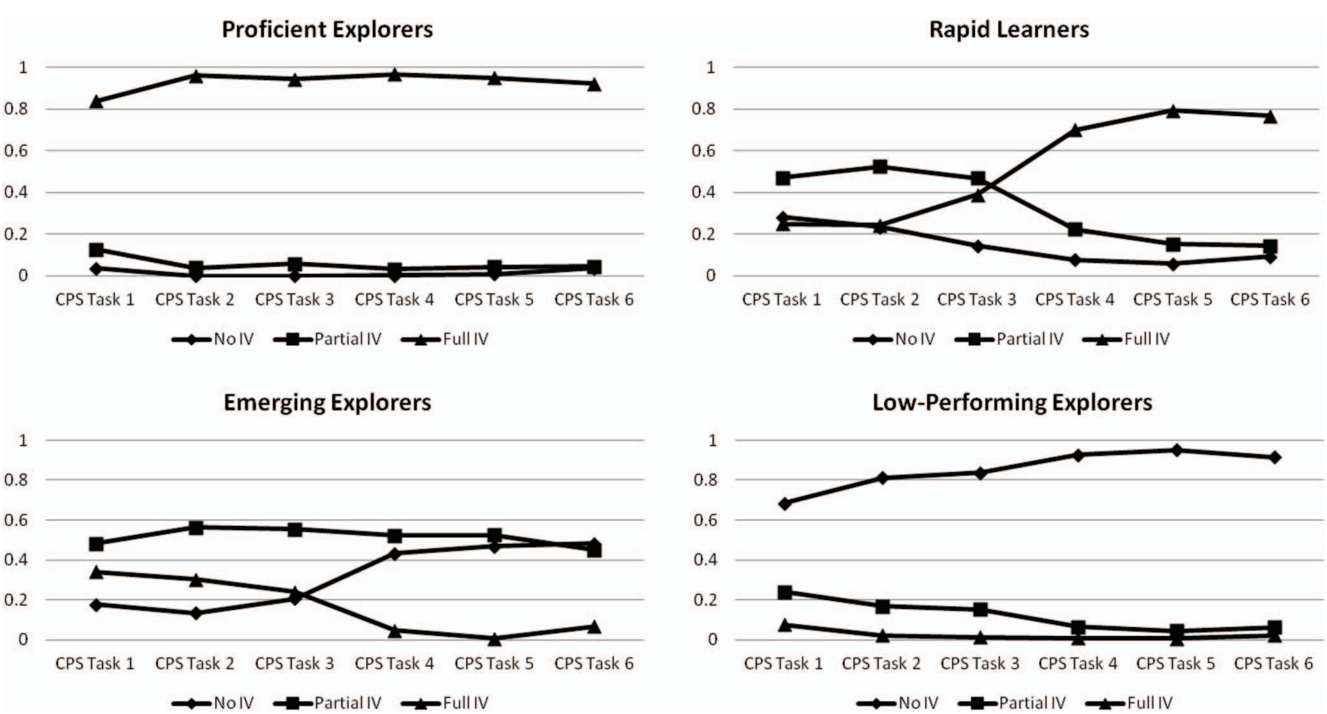

Figure 2. Item response probabilities that Proficient Explorers, Rapid Learners, Emerging Explorers, and Low-Performing Explorers would use the isolated variation strategy at Time 1 and Time 2.

Emerging Explorers, and *Low-Performing Explorers*) included $4^2 = 16$ possible developmental paths (i.e., transition patterns). The properties of these paths are listed in Table 4. We grouped these developmental paths, characterizing them as either *stable*, *improving*, or *declining*. Stable paths were the most frequently occurring developmental paths; adolescents on *stable paths* had been classified as having the same profile of isolated variation strategy use at Time 1 and Time 2 (43% of the adolescents). Adolescents on *improving paths* had been classified as having demonstrated more frequent use of the isolated variation strategy (i.e., as having a better strategy use profile) at Time 2 than Time 1 (36%). Finally, adolescents on *declining paths* had been classified as having demonstrated poorer use of the isolated variation strategy (i.e., as having a worse strategy use profile) at Time 2 than Time 1 (12%).

The latent transition probabilities indicate the conditional probability of having a particular strategy use profile at Time 2 (see Table 5). For instance, if adolescents were classified as *Proficient Explorers* or *Low-Performing Explorers* at Time 1, they were most likely to be classified as *Proficient Explorers* (.72) or *Low-Performing Explorers* (.59) at Time 2. In contrast, if adolescents were classified as *Rapid Learners* at Time 1, their performance was most likely to improve, resulting in their classification as *Proficient Explorers* by Time 2 (.55). Adolescents who were classified as *Emerging Explorers* at Time 1 were most likely to be classified as *Low-Performing Explorers* at Time 2 (.37).

Taken together, the profiles of the greatest number of adolescents remained stable: Adolescents who were classified as *Proficient Explorers* or *Low-Performing Explorers* at Time 1 had a high probability of retaining the same profile. However, there was also a smaller number of adolescents whose profiles changed: *Rapid Learners* had a relatively high probability of achieving a better profile at Time 2, and *Emerging Explorers* had a relatively high risk of having a worse profile at Time 2.

Table 5
Transition Probabilities of Profiles From Time 1 (Columns) to Time 2 (Rows)

Profile	*Proficient Explorers*	*Rapid Learners*	*Emerging Explorers*	*Low Performing Explorers*
Proficient Explorers	.72	.10	.05	.13
Rapid Learners	.55	.14	.06	.24*
Emerging Explorers	.26	.19**	.18**	.37**
Low Performing Explorers	.19	.08*	.14**	.59**

Note. The group of *Proficient Explorers* serves as the reference profile in the multinomial regression.
* $p < .05$. ** $p < .001$.

Goal 3: Test the Impact of Reasoning Abilities and Learning-Related Motivational Beliefs on Transition Probability

Finally, we incorporated reasoning and learning-related motivational beliefs as separate covariates into the LTA model to predict the adolescents' latent class transition probabilities. All models with covariates had a lower AIC and BIC and higher entropy than models without covariates.

Neither scientific reasoning nor figural reasoning abilities predicted adolescents' strategy use profiles at Time 1 or their patterns of transition from Time 1 to Time 2 profiles. However, adolescents' verbal reasoning scores predicted the probability that those classified as *Emerging Explorers* at Time 1 were classified as *Low Performing Explorers* at Time 2 ($B = -1.01$, $SE = 0.42$, $p = .016$, $OR = 0.36$): The higher *Emerging Explorers*' verbal reasoning scores, the less likely they were to be classified as *Low Performing Explorers* (as compared with *Proficient Explorers*) at Time 2.

Learning-related motivational beliefs did not predict other transition patterns, but they did predict the probability that those classified as *Emerging Explorers* at Time 1 would also be classi-

Table 4
Absolute and Relative Frequencies of the Sample in the Developmental Paths of Complex Problem-Solving Behavior

Path	Time 1	Time 2	Absolute number	Sample proportion
Stable	*Low Performing Explorers*	*Low Performing Explorers*	416	.21
Stable	*Proficient Explorers*	*Proficient Explorers*	310	.15
Declining	*Emerging Explorers*	*Low Performing Explorers*	198	.10
Improving	*Rapid Learners*	*Proficient Explorers*	180	.09
Improving	*Emerging Explorers*	*Proficient Explorers*	142	.07
Improving	*Low Performing Explorers*	*Proficient Explorers*	135	.07
Improving	*Emerging Explorers*	*Rapid Learners*	104	.05
Improving	*Low Performing Explorers*	*Emerging Explorers*	100	.05
Stable	*Emerging Explorers*	*Emerging Explorers*	94	.05
Declining	*Rapid Learners*	*Low Performing Explorers*	78	.04
Improving	*Low Performing Explorers*	*Rapid Learners*	60	.03
Declining	*Proficient Explorers*	*Low Performing Explorers*	58	.03
Stable	*Rapid Learners*	*Rapid Learners*	47	.02
Declining	*Proficient Explorers*	*Rapid Learners*	41	.02
Declining	*Proficient Explorers*	*Emerging Explorers*	22	.01
Declining	*Rapid Learners*	*Emerging Explorers*	21	.01

fied at Time 2 ($B = -0.70$, $SE = 0.26$, $p = .007$, $OR = 0.50$): The more positive learning-related motivational beliefs *Emerging Explorers* held, the less likely it was for them to be reclassified as *Emerging Explorers* (as compared with *Proficient Learners*) at Time 2.

Discussion

The three main goals of the present study were to (a) determine the optimal number of strategy use profiles in complex problem solving of adolescents, (b) determine the number of meaningful developmental paths for each strategy use profile, and (c) test the impact of reasoning abilities and learning-related motivational beliefs on the probability that an adolescent with a given strategy use profile will take a given developmental path. The present study showed that LTA is a valuable tool to investigate and reveal developmental paths of adolescents' strategy use profiles.

First, four meaningful profiles emerged from the latent transition analysis: *Proficient Explorers*, *Rapid Learners*, *Emerging Explorers*, and *Low-Performing Explorers*. These results are in line with the results obtained in a large Hungarian sample using a cross-sectional design (Greiff et al., 2018). Replication is an important research endeavor and this study met this endeavor by replicating strategic behavior profiles in an adolescent sample in Finland using a longitudinal design, thereby showing that the findings are generalizable across countries (Duncan et al., 2007).

Second, we found three developmental pathways (*stable*, *improving*, and *declining*) and showed that some adolescent profiles were more stable over time than others. We found that most of the adolescents in our sample showed stable developmental pathways (i.e., no change in their use of the isolated variation strategy): Adolescents classified as *Proficient Explorers* or *Low-Performing Explorers* at Time 1 were most likely to be classified at Time 2. The CPS tasks clearly separated these two extreme profiles of adolescents, whose development was less influenced by reasoning ability and learning-related motivational beliefs than the profiles of *Rapid Learners* and *Emerging Explorers* were. On the other hand, adolescents classified as *Rapid Learners* and *Emerging Explorers* at Time 1 were more likely to be reclassified at Time 2. *Rapid Learners'* strategy use was most likely to improve, resulting in their classification as *Proficient Explorers* at Time 2, and *Emerging Explorers'* strategy use was most likely to decline, resulting in their classification as *Low-Performing Explorers* at Time 2.

Third, we found that *Emerging Explorers'* strategy use was less likely to decline at Time 2 if their verbal reasoning scores were high. Moreover, *Emerging Explorers* with positive learning-related motivational beliefs showed improved strategy use at Time 2. These findings may indicate that, although CPS task performance is rather independent of prior knowledge, a certain degree of verbal reasoning ability enables adolescents to perform well on the task. CPS tasks entail particular verbal components that require some degree of verbal reasoning to understand the task requirements, at least in the initial stages of the task. PISA data confirm that CPS scores are associated with achievements in mathematics, science *and* reading (Wirth, Leutner, & Klieme, 2005); therefore, better verbal reasoning skills have likely protected the strategy use of *Emerging Explorers* against decline over time. These results need further empirical support because the reliability of the verbal reasoning test is lower than the reliability of the other measures. However, the results indicate that the positive development of CPS strategy use in particular groups of adolescents is associated with better verbal reasoning abilities.

Most important for the present study is that learning-related motivational beliefs, which were assessed on a general level, had specific effects on the adolescents' strategy use over time. Positive learning-related motivational beliefs facilitated *Emerging Explorers'* improvement in strategy use and, thus, their classification as *Proficient Explorers* at Time 2. This finding supports the common notions that one's perception of oneself as capable of learning is associated with a positive learning effect. The results on the effect of verbal ability and learning-related motivational beliefs confirm previous longitudinal findings (Vainikainen et al., 2015) in which verbal ability and learning-related beliefs were shown to maintain and facilitate adolescents' CPS development. Learning-related motivational beliefs and verbal abilities might be a good starting point for interventions to promote the favorable development of strategic behavior during CPS in the group of *Emerging Explorers*.

The finding that none of the three reasoning abilities or the learning-related motivational beliefs influenced the positive development of the *Rapid Learners'* strategic profile was unexpected. A cross-sectional large-scale study found fluid reasoning, scientific reasoning, and learning orientation to be good predictors of the isolated variation strategy use during CPS, but also showed that all predictors together still left about 60% of the variance in use of the isolated variation strategy unexplained (Wüstenberg et al., 2014). The authors speculated that metastrategic knowledge may be a relevant, but not investigated influence on the use of the strategy. We were not able to test this in this study, but following this rationale, *Rapid Learners* might have acquired some metastrategic knowledge between T1 and T2 that helped them to become *Proficient Explorers* at T2. Factual knowledge about strategies, reasoning, and motivational beliefs, might, in concert with the adolescents' personal background, developmental history, and persistence, better explain the development of strategy use during CPS. Future studies should examine a range of other important individual and social-contextual factors (e.g., persistence or sociocultural background) together with measures of metastrategic knowledge to complement current findings on the development of *Rapid Learners* or *Low-Perfoming Explorers*.

Overall, the observed differences between students are consistent with the cognitive load theory (van Merriënboer & Sweller, 2005). There is considerable variability in cognitive effort required to master complex learning material (here: the CPS tasks). Thus, the differences pertain to the level of cognitive load because students' access to the relevant cognitive schemata that reflect how to explore the CPS tasks is available to different degrees (Greiff et al., 2018). This view is helpful as it links our findings to the vast body of literature on instructional design and intervention research embedded in cognitive load research (e.g., Paas, Renkl, & Sweller, 2004). So far, interventions aimed at increasing CPS performance have relied on repeatedly confronting problem-solvers with problems of a similar nature (e.g., Kretzschmar & Süß, 2015). Training lead to an increase in performance and also to an increase in strategic behavior (Lotz et al., 2017). However, to the best of our knowledge, no efforts have been dedicated toward instructional design trying to understand different profiles of CPS performance. Based on our findings, trainings in CPS strategic behavior could be more specifically tailored to the needs of specific groups of students providing them with the scaffolding necessary to succeed in solving complex problems.

The current findings also confirm that the assumptions of the overlapping wave model may be applied on CPS strategy use development (Siegler, 1996). In line with the overlapping waves model, we showed considerable within and between differences between students' CPS strategy use over time. Also, we found various developmental paths, for example, groups of students who progressed to more advanced CPS strategy use profiles as well as groups who shifted to worse profiles over time. Particularly *Emerging Explorers*, who used the isolated variation strategy inconsistently, were likely to decline to a worse strategy use profile before being able to apply the best strategy for the task.

It is important to note that our results only apply to 11- to 17-year-old adolescents and one specific measure of CPS. The findings appear to be generalizable across Europe (Hungary and Finland; Greiff et al., 2018); future studies need to confirm whether the findings generalize to other age groups, other cultures, and other CPS measures (e.g., Genetics Lab: Sonnleitner et al., 2012; MultiFlux: Kröner, Plass, & Leutner, 2005).

Also, while using the LTA method, we carefully considered statistical, theoretical, and parsimony criteria for interpreting the profiles. However, it is important to note that, just as in all other scientific studies using LTA or LCA, the meaning of the profiles was assigned post hoc as neither method is confirmatory. Finally, because of the 3 years' time between Time 1 and Time 2, it is impossible to know if the adolescents transitioned directly from one profile to another from Time 1 to Time 2. One would need to consider the underlying speed and progress of the behavior change in each profile to reach conclusions about what happened during the 3 years. Unfortunately, the present design does not provide these insights. Future studies are needed to complement the findings using either shorter time frames between measurement time points or implementing intensive longitudinal designs, such as ambulatory assessments (Fahrenberg, Myrtek, Pawlik, & Perrez, 2007).

The investigation of continuity and change in behavior is a long-standing developmental science interest (Bornstein, Putnick, & Esposito, 2017). To our knowledge, this is the first longitudinal study to explore the continuity of and change in developmental pathways of adolescents' strategic behavior in complex environments and antecedents of that behavior in a large sample of adolescents using a latent mixture modeling approach. We identified four profiles of adolescents following either stable, improving, or declining pathways. Verbal reasoning ability and learning-oriented motivational beliefs were associated with a positive development of *Emerging Explorers*, a group of adolescents that tends to occasionally use the isolated variation strategy over time. Verbal reasoning and learning oriented motivational beliefs served as protective and reinforcing factors in the development of these adolescents.

The identification of profiles as well as their respective developmental pathways and relevant predictors are important because they allow researchers and practitioners to detect individual adolescents' needs and to tailor future adolescents' education and training. Because of the increasing relevance of domain-general skills such as CPS, the continued pursuit of this research and training agenda will have significant implications for educational and developmental psychology.

References

Agresti, A. (2002). *Categorical data analysis*. New York, NY: Wiley. http://dx.doi.org/10.1002/0471249688

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52,* 317–332. http://dx.doi.org/10.1007/BF02294359

Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science, 1,* 164–180. http://dx.doi.org/10.1111/j.1745-6916.2006.00011.x

Bornstein, M. H., Hahn, C.-S., Bell, C., Haynes, O. M., Slater, A., Golding, J., . . . the ALSPAC Study Team. (2006). Stability in cognition across early childhood. A developmental cascade. *Psychological Science, 17,* 151–158. http://dx.doi.org/10.1111/j.1467-9280.2006.01678.x

Bornstein, M. H., Putnick, D. L., & Esposito, G. (2017). Continuity and stability in development. *Child Development Perspectives, 11,* 113–119. http://dx.doi.org/10.1111/cdep.12221

Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*. San Francisco, CA: Holden-Day.

Buchner, A. (1995). Basic topics and approaches to the study of complex problem solving. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 27–63). Hillsdale, NJ: Erlbaum.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development, 70,* 1098–1120. http://dx.doi.org/10.1111/1467-8624.00081

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Hoboken, NJ: Wiley.

Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6,* 330–351. http://dx.doi.org/10.1037/1082-989X.6.4.330

Danner, D., Hagemann, D., Schankin, A., Hager, M., & Funke, J. (2011). Beyond IQ: A latent state-trait analysis of general intelligence, dynamic decision making, and implicit learning. *Intelligence, 39,* 323–334. http://dx.doi.org/10.1016/j.intell.2011.06.004

Dörner, D., & Güss, C. D. (2013). PSI: A computational architecture of cognition, motivation, and emotion. *Review of General Psychology, 17,* 297–317. http://dx.doi.org/10.1037/a0032947

Dunbar, K., & Klahr, D. (2012). Scientific thinking and reasoning. In K. J. Holoyak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 701–718). New York, NY: Oxford University Press.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43,* 1428–1446. http://dx.doi.org/10.1037/0012-1649.43.6.1428

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41,* 1040–1048. http://dx.doi.org/10.1037/0003-066X.41.10.1040

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53,* 109–132. http://dx.doi.org/10.1146/annurev.psych.53.100901.135153

Fahrenberg, J., Myrtek, M., Pawlik, K., & Perrez, M. (2007). Ambulatory assessment-monitoring behavior in daily life settings. *European Journal of Psychological Assessment, 23,* 206–213. http://dx.doi.org/10.1027/1015-5759.23.4.206

Feldman, B. J., Masyn, K. E., & Conger, R. D. (2009). New approaches to studying problem behaviors: A comparison of methods for modeling longitudinal, categorical adolescent drinking data. *Developmental Psychology, 45,* 652–676. http://dx.doi.org/10.1037/a0014851

Findley, M. J., & Cooper, H. M. (1983). Locus of control and academic achievement: A literature review. *Journal of Personality and Social Psychology, 44,* 419–427. http://dx.doi.org/10.1037/0022-3514.44.2.419

Frischkorn, G. T., Greiff, S., & Wüstenberg, S. (2014). The development of complex problem solving in adolescence: A latent growth curve analysis. *Journal of Educational Psychology, 106,* 1007–1020. http://dx.doi.org/10.1037/a0037114

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science, 7,* 237–241. http://dx.doi.org/10.1111/j.1467-9280.1996.tb00366.x

Gagné, F., & St. Père, F. (2001). When IQ is controlled, does motivation still predict achievement? *Intelligence, 30,* 71–100. http://dx.doi.org/10.1016/S0160-2896(01)00068-X

Greiff, S., Molnár, G., Martin, R., Zimmermann, J., & Csapó, B. (2018). Students' exploration strategies in computer-simulated complex problem environments: A latent class approach. *Computers & Education, 125,* 248–263. http://dx.doi.org/10.1016/j.compedu.2018.07.013

Greiff, S., Stadler, M., Sonnleitner, P., Wolff, C., & Martin, R. (2015). Sometimes less is more: Comparing the validity of complex problem-solving measures. *Intelligence, 50,* 100–113. http://dx.doi.org/10.1016/j.intell.2015.02.007

Greiff, S., & Wüstenberg, S. (2014). Assessment with microworlds using MicroDYN: Measurement invariance and latent mean comparisons. *European Journal of Psychological Assessment, 30,* 304–314. http://dx.doi.org/10.1027/1015-5759/a000194

Greiff, S., Wüstenberg, S., & Avvisati, F. (2015). Computer-generated log-file analyses as a window into students' minds? A showcase study based on the PISA 2012 assessment of problem solving. *Computers & Education, 91,* 92–105. http://dx.doi.org/10.1016/j.compedu.2015.10.018

Greiff, S., Wüstenberg, S., & Funke, J. (2012). Dynamic problem solving: A new measurement perspective. *Applied Psychological Measurement, 36,* 189–213. http://dx.doi.org/10.1177/0146621612439620

Greiff, S., Wüstenberg, S., Goetz, T., Vainikainen, M.-P., Hautamäki, J., & Bornstein, M. H. (2015). A longitudinal study of higher-order thinking skills: Working memory and fluid reasoning in childhood enhance complex problem solving in adolescence. *Frontiers in Psychology, 6,* 1060. http://dx.doi.org/10.3389/fpsyg.2015.01060

Hautamäki, J. (1984). *Peruskoululaisten loogisen ajattelun mittaamisesta ja esiintymisestä* [Measuring and the occurrence of logical thinking among basic school adolescents]. Joensuu, Finland: Social Sciences Publications of the University of Joensuu.

Hautamäki, J., Arinen, P., Eronen, S., Hautamäki, A., Kupiainen, S., Lindblom, B., . . . Scheinin, P. (2002). *Assessing learning-to-learn: A framework*. Helsinki, Finland: National Board of Education.

Hipp, J. R., & Bauer, D. J. (2006). Local solutions in the estimation of growth mixture models. *Psychological Methods, 11,* 36–53. http://dx.doi.org/10.1037/1082-989X.11.1.36

Hulleman, C. S., Schrager, S. M., Bodmann, S. M., & Harackiewicz, J. M. (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? *Psychological Bulletin, 136,* 422–449. http://dx.doi.org/10.1037/a0018947

Kam, C., Morin, A. J. S., Meyer, J. P., & Topolnytsky, L. (2016). Are commitment profiles stable and predictable? A latent transition analysis. *Journal of Management, 42,* 1462–1490. http://dx.doi.org/10.1177/0149206313503010

Kretzschmar, A., & Süß, H. M. (2015). A study on the training of complex problem solving competence. *Journal of Dynamic Decision Making, 1,* 1–14.

Kröner, S., Plass, J. L., & Leutner, D. (2005). Intelligence assessment with computer simulations. *Intelligence, 33,* 347–368. http://dx.doi.org/10.1016/j.intell.2005.03.002

Kuhn, D. (2010). What is scientific thinking and how does it develop? In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (pp. 371–393). Oxford, United Kingdom: Wiley-Blackwell. http://dx.doi.org/10.1002/9781444325485.ch19

Kuhn, D., Garcia-Mila, M., Zohar, A., Andersen, C., White, S. H., Klahr, D., & Carver, S. M. (1995). Strategies of knowledge acquisition. *Monographs of the Society for Research in Child Development, 60,* 1–157. http://dx.doi.org/10.2307/1166059

Kyllonen, P., Anguiano Carrasco, C., & Kell, H. J. (2017). Fluid ability (Gf) and complex problem solving (CPS). *Journal of Intelligence, 5,* 28. http://dx.doi.org/10.3390/jintelligence5030028

Lanza, S. T., & Collins, L. M. (2008). A new SAS procedure for latent transition analysis: Transitions in dating and sexual risk behavior. *Developmental Psychology, 44,* 446–456. http://dx.doi.org/10.1037/0012-1649.44.2.446

Lanza, S. T., Patrick, M. E., & Maggs, J. L. (2010). Latent transition analysis: Benefits of a latent variable approach to modeling transitions in substance use. *Journal of Drug Issues, 40,* 93–120. http://dx.doi.org/10.1177/002204261004000106

Lotz, C., Scherer, R., Greiff, S., & Sparfeldt, J. R. (2017). Intelligence in action: Effective strategic behaviors while solving complex problems. *Intelligence, 64,* 98–112. http://dx.doi.org/10.1016/j.intell.2017.08.002

Molnár, G., Greiff, S., & Csapó, B. (2013). Inductive reasoning, domain specific and complex problem solving: Relations and development. *Thinking Skills and Creativity, 9,* 35–45. http://dx.doi.org/10.1016/j.tsc.2013.03.002

Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Author.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., . . . Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51,* 77–101. http://dx.doi.org/10.1037/0003-066X.51.2.77

Nettelbeck, T., & Burns, N. R. (2010). Processing speed, working memory and reasoning ability from childhood to old age. *Personality and Individual Differences, 48,* 379–384. http://dx.doi.org/10.1016/j.paid.2009.10.032

OECD. (2014). *PISA 2012 results: Creative problem solving: Adolescents' skills in tackling real-life problems* (Vol. V). Paris, France: OECD Publishing.

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science, 32,* 1–8. http://dx.doi.org/10.1023/B:TRUC.0000021806.17516.d0

Piaget, J., & Inhelder, B. (1956). *The child's conception of space* (F. J. Langdon & J. L. Lunzer, Trans.). London, United Kingdom: Routledge & Kegan Paul.

Polk, T. A., & Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review, 102,* 533–566. http://dx.doi.org/10.1037/0033-295X.102.3.533

Raven, J. (2000). The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology, 41,* 1–48. http://dx.doi.org/10.1006/cogp.1999.0735

Rawsthorne, L. J., & Elliot, A. J. (1999). Achievement goals and intrinsic motivation: A meta-analytic review. *Personality and Social Psychology Review, 3,* 326–344. http://dx.doi.org/10.1207/s15327957pspr0304_3

Rinne, L. F., Ye, A., & Jordan, N. C. (2017). Development of fraction comparison strategies: A latent transition analysis. *Developmental Psychology, 53,* 713–730. http://dx.doi.org/10.1037/dev0000275

Ross, J. D., & Ross, C. M. (1979). *Ross test of higher cognitive processes*. Novato, CA: Academic Therapy Publications.

Rotter, J. B. (1990). Internal versus external control of reinforcement: A case history of a variable. *American Psychologist, 45,* 489–493. http://dx.doi.org/10.1037/0003-066X.45.4.489

Rudolph, J., Niepel, C., Greiff, S., Goldhammer, F., & Kröner, S. (2017). Metacognitive confidence judgements and their link to complex problem solving. *Intelligence, 63,* 1–8. http://dx.doi.org/10.1016/j.intell.2017.04.005

Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics, 6,* 461–464. http://dx.doi.org/10.1214/aos/1176344136

Schweizer, F., Wüstenberg, S., & Greiff, S. (2013). Validity of the MicroDYN approach: Complex problem solving predicts school grades beyond working memory capacity. *Learning and Individual Differences, 24,* 42–52. http://dx.doi.org/10.1016/j.lindif.2012.12.011

Shayer, M. (1979). Has Piaget's construct of formal operational thinking any utility? *British Journal of Educational Psychology, 49,* 265–276. http://dx.doi.org/10.1111/j.2044-8279.1979.tb02425.x

Siegler, R. (1996). *Emerging minds: The process of change in children's thinking*. Oxford: Oxford University Press.

Sonnleitner, P., Brunner, M., Greiff, S., Funke, J., Keller, U., Martin, R., . . . Latour, T. (2012). The Genetics Lab: Acceptance and psychometric characteristics of a computer-based microworld assessing complex problem solving. *Psychological Test and Assessment Modeling, 54,* 54–72.

Stadler, M., Niepel, C., & Greiff, S. (2019). Differentiating between static and complex problems: A theoretical framework and its empirical validation. *Intelligence, 72,* 1–12. http://dx.doi.org/10.1016/j.intell.2018.11.003

Tschirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development, 51,* 1–10. http://dx.doi.org/10.2307/1129583

Utman, C. H. (1997). Performance effects of motivational state: A meta-analysis. *Personality and Social Psychology Review, 1,* 170–182. http://dx.doi.org/10.1207/s15327957pspr0102_4

Vainikainen, M.-P. (2014). *Finnish primary school pupils' performance in learning to learn assessments: A longitudinal perspective on educational equity*. Helsinki, Finland: Helsingin yliopisto.

Vainikainen, M.-P., Wüstenberg, S., Kupiainen, S., Hotulainen, R., & Hautamäki, J. (2015). Development of learning to learn skills in primary school. *International Journal of Lifelong Education, 34,* 376–392. http://dx.doi.org/10.1080/02601370.2015.1060025

van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning. Recent developments and future directions. *Educational Psychology Review, 17,* 147–177. http://dx.doi.org/10.1007/s10648-005-3951-0

Wirth, J., Leutner, D., & Klieme, E. (2005). Problemlosekompetenz: Okonomisch und zugleich differenziert erfassbar? [Problem-solving competence: Economically assessable and differentiable at the same time?] In E. Klieme, D. Leutner, & J. Wirth (Eds.), *Students' problem-solving competence: Diagnostic approaches, theoretical foundations, and empirical findings of the German PISA-2000 study* (pp. 73–82). Wiesbaden, Germany: Verlag für Sozialwissenschaften. http://dx.doi.org/10.1007/978-3-322-85144-4_6

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving: More than reasoning? *Intelligence, 40,* 1–14. http://dx.doi.org/10.1016/j.intell.2011.11.003

Wüstenberg, S., Stadler, M., Hautamäki, J., & Greiff, S. (2014). The role of strategy knowledge for the application of strategies in complex problem-solving tasks. *Technology, Knowledge and Learning, 19,* 127–146. http://dx.doi.org/10.1007/s10758-014-9222-8

Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review, 20,* 99–149. http://dx.doi.org/10.1006/drev.1999.0497

Appendix

The LTA Mathematical Model

Suppose a latent transition model with n_s latent statuses is to be estimated based on a data set including M categorical items measured at each of T times for a total of MT items, a covariate X, and a grouping variable G. Let $\mathbf{Y}_i = (Y_{i11}, Y_{i12}, \ldots, Y_{i1M}, Y_{i21}, Y_{i22}, \ldots, Y_{i2M}, Y_{iT1}, Y_{iT2}, \ldots, Y_{iTM})$ represent the vector of individual i's responses for all times $t = 1, \ldots, T$ and items $m = 1, \ldots, M$, where an individual response Y_{itm} may take on the values $1, 2, \ldots, r_m$. Let $s_{1i} = 1, 2, \ldots, n_s$ be individual i's latent status membership at Time 1, $s_{2i} = 1, 2, \ldots, n_s$ be individual i's latent status membership at Time 2, and so on. Let $I\ (y = k)$ be the indicator function which equals 1 if response y equals k and 0 otherwise. Suppose also that G_i represents the value of individual i's group membership, X_i represents the value of the covariate X for individual i and that the value of X can relate to the probability of membership in each latent status, δ, and each transition probability, τ. Then the latent transition model can be expressed as:

$$\mathrm{P}(Y_i = y \mid X_i = x, G_i = g)$$
$$= \sum_{s_1=1}^{n_s} \cdots \sum_{s_1=1}^{n_s} \delta_{s_1|g}(x)\tau_{s_1|s_2,g}(x) \cdots \tau_{s_t|s_{t-1},g}(x) \prod_{m=1}^{M} \prod_{k=1}^{r_m} \prod_{t=1}^{T} \rho_{mk|s_t,g}^{I(y_m=k)}$$

where $\delta_{s1|g}(x) = P(S_{1i} = s_i \mid X_i = x, G_i = g)$ is a standard baseline-category multinomial logistic model (Agresti, 2002) predicting individual i's membership latent status s_1 at Time 1.

Note. Reprinted from "A new SAS procedure for latent transition analysis: Transitions in dating and sexual risk behavior," by S. T. Lanza and L. M. Collins (2008), *Developmental Psychology, 44,* 449. Copyright 2008 by the American Psychological Association.

Received February 28, 2018
Revision received April 11, 2019
Accepted May 15, 2019 ■