

The logic of success: the relation between complex problem-solving skills and university achievement

Matthias Stadler¹ · Nicolas Becker² · Johannes Schult² ·
Christoph Niepel¹ · Frank M. Spinath² ·
Jörn R. Sparfeldt² · Samuel Greiff¹

Published online: 2 September 2017
© Springer Science+Business Media B.V. 2017

Abstract The successful completion of a university degree program is accompanied by multiple complex opportunities and challenges, which require students to react accordingly with the skills necessary to meet them. Therefore, the aim of this study was to investigate the role of complex-problem solving (CPS) skills in undergraduate students' university success in two independent samples. In Study 1, 165 university students completed a measure of intelligence as well as a measure of CPS. In addition, students' university grade point averages (GPAs) and their subjective evaluation of academic success were collected. CPS made a significant contribution to the explanation of GPAs and the subjective success evaluations even when controlling for intelligence. To further investigate this effect, Study 2 relied on an independent and more heterogeneous sample of 216 university students. The findings of Study 1 were essentially replicated in Study 2. Thus, the combined results demonstrate a link between individual differences in CPS and the abilities necessary to be academically successful in university education.

Keywords University success · GPA · Complex problem solving · Intelligence · Cognitive ability · Structural equation modeling

Attending a university is becoming more and more commonplace in modern societies (Pittman and Richmond 2008), with an increasing number of students enrolling in university programs and societies investing large amounts of money in their educational systems (OECD 2014). For the individual students, the transition from high school to university life constitutes a

✉ Matthias Stadler
matthias.stadler@ur.de

¹ University of Luxembourg, ECCS, Maison des Sciences Humaines, 11, Porte des Sciences, 4366 Esch-sur-Alzette, Luxembourg

² Saarland University, Saarbrücken, Germany

critical life event (e.g., Terenzini et al. 1994) with unique opportunities as well as challenges: opportunities, because the scope of independent exploration of life's possibilities is greater than it will be at any other period of the life course; challenges, because independence and autonomy can also imply disorientation and uncertainty (Arnett 2000).

Researchers are thus increasingly interested in identifying and examining factors which are related to how students navigate successfully through their university years (Tavernier and Willoughby 2014). This has resulted in a vast array of cognitive, noncognitive, and demographic factors known to influence students' university success (e.g., Richardson et al. 2012). Within this study, we will focus primarily on individual differences in cognitive variables related to students' university performance. Most prominently, previous academic achievement and intelligence have been established as valid predictors of students' grade point average (GPA; Formazin et al. 2011). As an addition to these established cognitive predictors, academic-related skills such as problem solving were found to be important antecedents of students' success at university (Robbins et al. 2004). In that, individual differences in complex problem-solving (CPS) skills, that is, the skills necessary to deal with new and dynamically changing situations (Frensch and Funke 1995), might provide valuable information in explaining why students succeed differently well at university. Recent research has provided first evidence that CPS is significantly related to academic performance in school (Wüstenberg et al. 2012; Greiff et al. 2013) and at university (Stadler et al. 2015a) with incremental validity over and above intelligence. The present two-study report aims at expanding upon this initial evidence by investigating the relation between university students' skills at dealing with complex problems and their success at university in two studies with independent samples.

Measuring university success

Understanding university success depends on being able to conceptually define as well as assess it in a reliable and valid way (Richardson et al. 2012). Students' academic performance is usually expressed in terms of GPA representing the mean of the grades received in courses contributing to the final degree (Richardson et al. 2012). GPA is the most widely used and studied measure in tertiary education (Bacon and Bean 2006; Richardson et al. 2012), is economically available, shows good internal reliability and temporal stability (e.g., Bacon and Bean 2006), and correlates strongly with variables of interest to educational researchers such as intelligence, motivational strategies, or certain personality traits (Richardson et al. 2012). GPA represents a key criterion for postgraduate selection and employment and has been found to be a valid predictor of socioeconomic success (Strenze 2007). Moreover, GPA shows very strong correlations to other indicators of university success such as retention (Robbins et al. 2004). As such, it is an index that is directly meaningful to students, universities, and employers alike and relevant to future training and employment opportunities.

Nonetheless, the use of GPA as an indicator of university success has often been criticized (e.g., Johnson 2003). Grading disparities between universities, study programs, and even between different examiners impair a fair, reliable, and valid assessment of students' competencies. This has serious consequences on their future perspectives with respect to completing their university education with a higher GPA and, thus, better career prospects.

Beyond a narrow focus on GPA, university success can furthermore be defined as a multidimensional construct with substantial subjective components such as individual perceptions of accomplishment, comparisons to peers, or future prospects (Gattiker and Larwood

1986). Solely focusing on objective criteria such as GPA does not encompass this intrinsic and subjective aspect of success and thus ignores an important dimension of how students evaluate their performance. Hughes (1959) pointed out the importance of which beholder's eye is viewing success. Subjective success is most pertinent from the perspective of the individual as individuals evaluate different facets of their career. Objective success is critical when considering an external perspective that "validates" the tangible facets of an individual's career. Correspondingly, researchers have argued that objective and subjective aspects of success should be considered complementary (Duckworth et al. 2012).

Predicting university success

Regardless of the specific conception of success, managing a university program requires dealing with a complex system of academic tasks, learning and study behaviors, social obligations, and various other demands that are dynamically changing and whose interrelations are not always obvious (Parker et al. 2004). Correspondingly, numerous cognitive (e.g., intelligence or previous academic achievement; Formazin et al. 2011), noncognitive (e.g., personality traits, motivational factors, self-regulatory learning strategies, or psychosocial contextual influences; for an overview see Richardson et al. 2012), and demographic (e.g., age or sociodemographic background; Robbins et al. 2004) factors have been established to influence students' university success.

Apart from intelligence, one of the strongest predictors of academic achievement (e.g., Jensen 1998; Kuncel, Hezlett, & Ones, 2001), other cognitive abilities have been in the focus of researchers recently. Especially in tertiary education, student selection procedures reduce variation in intelligence scores (Furnham et al. 2003). This is particularly important for universities as highly selective academic institutions (Jensen 1998). Consequently, factors other than intelligence may add important incremental information to the accurate prediction of university performance.

Substantial differences in the development and prediction of GPA and subjective indicators of university success have been reported (e.g., Harackiewicz et al. 2002). Whereas cognitive ability consistently predicts university students' GPA, subjective indicators of university success seem to be more closely linked to psychosocial and study skill factors (Robbins et al. 2006). In addition, CPS represents a more recently introduced academia-related concept. Recent research has provided initial evidence for the relevance of CPS for academic success at the university (see Stadler et al. 2015b). In this line of research, CPS can be defined 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." (Buchner, cited in Frensch and Funke 1995, p. 14.)

Being able to deal with dynamically changing and partially opaque systems is necessary to be successful at any academic institution. Support for this notion comes from several articles reporting CPS to predict high school grades even beyond measures of general intelligence (Greiff et al. 2013; Wüstenberg et al. 2012; see also Kretzschmar et al. 2016; Lotz et al. 2016). Compared to high school, the demands posed by university programs should be even more complex and cognitively challenging. In her model of university success, Ferrett (2000)

describes cognitive skills such as time management, preparing for and taking examinations, or using information resources as the focal point of the freshman year experience. In that, university students face a variety of new challenges such as learning and applying study habits in a more complex academic environment and generally discovering how to function as independent and academically successful adults, which requires planning and problem-solving competencies (e.g., acquiring knowledge about new problems or prioritizing subgoals). Surprisingly though, only one study has investigated the relation between CPS and university success to date. Stadler et al. (2015a) found a substantial relation between CPS and both GPA and subjective university success of business students ($\beta = .38$) that remained significant even after general intelligence was controlled for.

CPS, intelligence, and academic success

When investigating the relation between CPS and university success, it is important to consider the well-established association between CPS and intelligence (e.g., Funke and Frensch 2007; Wirth and Klieme 2003; Wüstenberg et al. 2012). On the one hand, intelligence measures are among the most consistently validated predictors of university success (Richardson et al. 2012), with average correlations of $r = .32$ between intelligence and GPA (corrected for attenuation; Hell et al. 2007). Intelligence thus possesses a high validity in the prediction of university success and shows incremental validity over high school GPA in predicting university GPA (Formazin et al. 2011). On the other hand, the conceptual and empirical relations of intelligence and CPS need to be considered. CPS and intelligence can theoretically be distinguished by the unique demands complex problems pose. There is, however, considerable theoretical overlap between CPS and intelligence as some characteristic features of CPS such as the integration of information are part of almost every definition of intelligence (Sternberg and Berg 1986). The majority of studies on the relation between CPS and intelligence correspondingly reports medium to strong correlations between the two constructs that were summarized in a meta-analysis reporting an average correlation of $\rho = .43$ between CPS and intelligence (Stadler et al. 2015b).

This study

Only Stadler et al. (2015a) have investigated the relation between CPS and university success to date. While their study provided first evidence supporting the role of CPS in university success, it was severely limited in its generalizability. First, the sample size used was rather small ($N = 78$) and did not allow for advanced statistical analyses such as structural equation modeling; second, the sample consisted exclusively of business students and was thus rather homogeneous and limited in generalizability. Finally, the type of CPS measure employed (FSYS; Wagener 2001) was shown to have unsatisfactory reliability (Greiff et al. 2015b).

This paper will therefore represent a necessary extension of Stadler et al.'s (2015a) study in various aspects investigating the role of CPS in university success. The aim of Study 1 is to expand upon the results reported by Stadler et al. (2015a) using a larger sample thus allowing for latent analyses on the construct level. Study 2 will go even further by investigating the relation between CPS and university success in a more heterogeneous sample. Moreover, considering exam scores as additional criteria and incorporating a longitudinal measurement might reveal further insights regarding the generalizability of these relations.

Study 1

Based on the findings presented above, Study 1 investigated the relation between CPS and students' university success. In line with the findings reported by Stadler et al. (2015a), we expected CPS to significantly predict both GPA and subjective indicators of university success (Hypothesis 1).

Despite the strong conceptual overlap between CPS and intelligence mentioned before, the relation between CPS and university success should not be solely due to a shared measurement of intelligence. Thus, we expected to find an incremental validity of CPS in predicting university students' success over intelligence (Hypothesis 2). However, there may be considerable differences in the strength of prediction of the two constructs. Intelligence is strongly linked to GPA but less strongly to subjective university success (e.g., Robbins et al. 2006). Correspondingly, the incremental validity of CPS should be stronger for subjective success than for GPA.

Method

Participants The sample consisted of 165 students recruited while attending lectures in the biology ($N = 46$), psychology ($N = 85$), and sports ($N = 34$) departments of a middle-sized German university. Sixty-one percent of the students were female, and the mean age was $M = 22.53$ years ($SD = 3.83$). The majority of students were in their fourth semester of studying at the university (equivalent to the second half of the sophomore year). All students attending the lectures participated in the assessment completely voluntarily.

Procedure All tests and questionnaires were conducted computer-based (lasting 90 min). Students were informed ahead that all personal data would be treated confidentially. Afterwards, they signed the informed consent sheet approved by the university's data protection agency.

Measures

GPA The GPA that had been achieved at the time of the study was employed as a relatively objective measure of university success. GPAs were retrieved from the official university sources (as approved by the students).

Subjective university success Consisting of five items, the scale to measure subjective university success (Stadler et al. 2015a) asked students to rate their agreement with statements such as "I am successful in my studies". Students responded on a Likert scale ranging from 1 to 5. The items were designed as an efficient way to assess subjective university success as a construct encompassing subjective evaluations of participants' success in relation to their effort, and in relation to their peers (see Gattiker and Larwood 1986). The scale showed good internal consistency ($\alpha = .80$).

CPS Individual differences in CPS abilities were assessed using 10 items based on the MicroDYN approach (Greiff et al. 2012; Greiff et al. 2015a, b). This set of items has been shown to provide highly reliable and valid CPS scores (e.g., Greiff et al. 2012). Figure 1

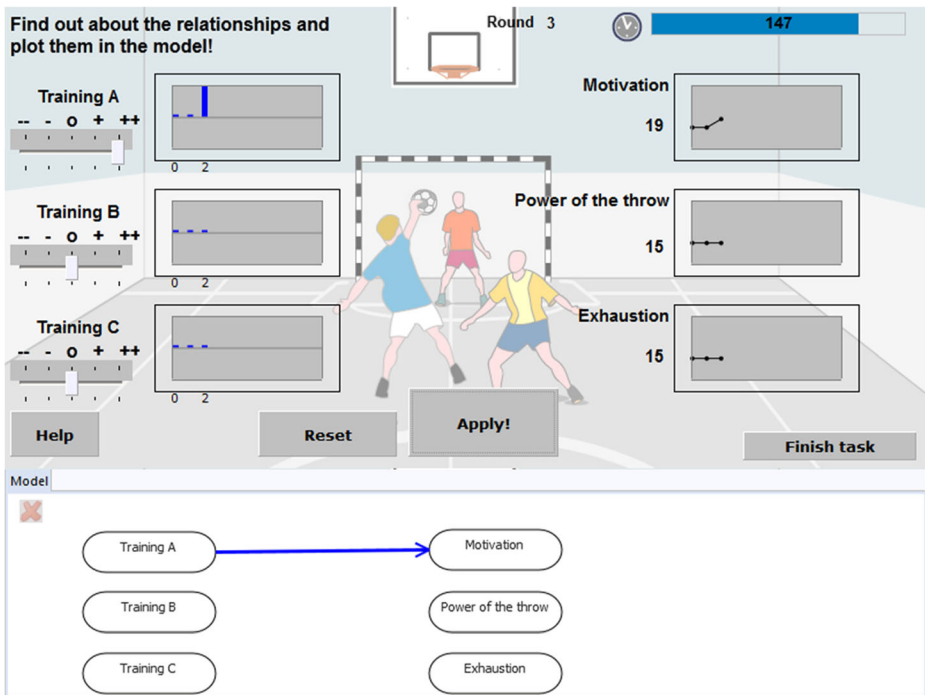


Fig. 1 Example of a CPS item based on the MicroDYN approach (Wüstenberg et al. 2012)

illustrates the type of MicroDYN tasks employed in our study. In this problem, the test taker is asked to explore the relation between three training strategies for handball players (labeled A, B, and C) and three outcomes (e.g., Motivation). The test taker may systematically vary the use of the training strategies to determine their effects on the outcomes. Unlike other measures of CPS emulating real-world problems, the underlying relations depicted here are arbitrary. At the end of the knowledge acquisition phase, participants were asked to plot the assumed relation at the bottom of the task. To reach certain predefined goals in the knowledge application phase (e.g., reach a Motivation value of 20 by adequately adapting the training methods), the acquired knowledge needed to be applied in a second step.

The scoring of students' CPS performance was conducted automatically by the testing software. For knowledge acquisition, credit (1 point) was given if the item's causal model was provided correctly; otherwise, no credit (0 points) was assigned. For knowledge application, credit (1 point) was given if all goals were reached in the application phase; otherwise, no credit was assigned (0 points).

Intelligence In order to determine participants' general intelligence, the Intelligenz-Struktur-Test-Screening (IST-Screening; Liepmann et al. 2012) was administered. The IST-Screening, as a well-established and economic (approximately 20 min) intelligence measure, consists of the three task groups of verbal analogies, number series, and figural matrices (each consisting of 20 items). The test's publishers report good internal consistencies for all three scales ($\alpha = .72\text{--}.90$).

Statistical analysis To test our hypotheses, we used structural equation modeling (SEM with WLSMV) in Mplus 7.3 (Muthén and Muthén 1998–2015). Model fit assessment was based on

fit indices recommended by Beauducel and Wittmann (2005) and the criteria proposed by Hu and Bentler (1999).

To account for the hierarchical structure of the data due to different university study programs, two dummy variables were created representing the three study programs with psychology students as the reference group. These dummy variables were added as control variables to all structural models to account for the within-cluster variance in a fixed effects model (Huang 2016).

Results

Descriptive statistics and measurement models Table 1 shows the descriptive statistics and observed intercorrelations for all variables included in Study 1. An ANOVA comparing the different study programs displayed significant differences in intelligence scores [$F(2;157) = 5.98; p = .003; \eta_p^2 = .084$], with psychology students ($M = 48.21; SD = 5.28$) averaging significantly higher scores than biology ($M = 44.57; SD = 6.56$) and sports students ($M = 43.97; SD = 9.03$).

To limit the parameters to be estimated in the structural models, we aggregated the intelligence items to three parcels for numerical, verbal, and figural content and had the parcels all load onto one factor. In line with previous research (e.g., Greiff and Neubert 2014), CPS was modeled as a higher-order factor consisting of the two latent factors of knowledge acquisition and knowledge application. Both for knowledge acquisition and knowledge application, all 10 items were defined to load onto one common factor each. All measurement models fitted very well to the data thus allowing for estimations of the structural models (Table 2).

Structural models In order to test Hypothesis 1, CPS was specified as a predictor of both subjective university success and GPA. This model represented the data very well (Table 2). In accordance with Hypothesis 1, CPS significantly predicted subjective university success ($\beta = .32; p = .004; R^2 = .10$) and GPA ($\beta = .34; p < .001; R^2 = .12$). The correlation between subjective university success and GPA was $r = .57 (p < .001)$.

To estimate the incremental validity of CPS over and above intelligence and to avoid issues of multicollinearity resulting from the high latent correlation between CPS and intelligence ($r = .81; p < .001$), CPS was residualized for intelligence in order to test Hypothesis 2. In this model, intelligence explained 66% ($\beta = .81; p < .001$) of the variance in CPS. The remaining residual of CPS, now not sharing any variance with intelligence, as well as intelligence itself were then defined to predict both subjective university success and GPA. In line with Hypothesis 2, the residual of CPS

Table 1 Descriptive statistics and observed correlations for Study 1

Variables	Mean (SD)	GPA	SUS	CPS acquisition	CPS application
GPA	2.31 (0.61)	–			
SUS	3.29 (0.59)	.33**	–		
CPS acquisition	0.52 (0.26)	.22**	.10	–	
CPS application	0.50 (0.23)	.19*	.23**	.70**	–
Intelligence	47.15 (2.95)	.08	.00	.24**	.27**

CPS complex problem solving, GPA grade point average, SUS subjective university success

* $p < .05$, ** $p < .01$

Table 2 Model fit indices for Study 1

Model	χ^2	df	<i>p</i>	CFI	TLI	RMSEA
Measurement models						
SUS ($\lambda = .45-.92$)	7.51	5	.19	.98	.99	.06
Intelligence ($\lambda = .51-.72$)	0	0	–	1.00	1.00	.00
CPS ($\lambda = .80-.93$)	207.81	168	.02	.97	.97	.04
Structural models						
CPS predicting GPA and SUS	339.63	296	.04	.97	.96	.03
CPS predicting GPA and SUS controlling for intelligence	409.26	371	.08	.97	.97	.03

df degrees of freedom, *CFI* comparative fit index, *TLI* Tucker-Lewis index, *RMSEA* root mean square error of approximation, *SUS* subjective university success, *CPS* complex problem solving, *GPA* grade point average

remained a significant predictor of both subjective university success ($\beta = .24; p < .001$) and GPA ($\beta = .14; p = .015$). As expected, CPS predicted subjective success significantly stronger than GPA after intelligence was controlled for ($\chi^2 = 2.77; df = 1; p = .048$). Intelligence itself also predicted subjective university success ($\beta = .23; p < .001$) and GPA ($\beta = .32; p < .001$). Intelligence predicted GPA significantly stronger than subjective university success ($\chi^2 = 8.43; df = 1; p = .002$). Together, intelligence and CPS explained 12% of the variance in GPA and 11% of the variance in subjective university success. The correlation between subjective university success and GPA was $r = .58$ ($p < .001$). As indicated by the fit indices, this model represents the data very well (Table 2).

Discussion of Study 1

The aim of Study 1 was to investigate the relation between CPS and students' university success. In line with our hypotheses and previous results (Stadler et al. 2015a), CPS was significantly related to both GPA and subjectively appraised success. This relation remained significant and substantial even after intelligence was controlled for.

Regarding the two different indicators of university success, there were considerable differences in the prediction by CPS and intelligence. CPS significantly predicted both GPA and subjective university success. In line with our hypotheses, the relation was significantly stronger between CPS and subjective university success than between CPS and GPA. Intelligence, on the other hand, also predicted both indicators of university success although it was more strongly related to GPA than to subjective university success. This confirms previous findings regarding differential prediction of GPA and subjective indicators of university success (e.g., Robbins et al. 2006) in that GPA is more strongly linked to intelligence while alternative indicators are related to other relevant skills.

In summary, the results of Study 1 support the validity of CPS as a predictor of university success, but they raise the question of whether the findings can be generalized to more heterogeneous samples as well or whether they only hold for specific, rather highly selective university programs. To further investigate the question of generalizability, Study 2 was based on a more heterogeneous sample.

Study 2

Study 2 expands on Study 1 by replicating its design as closely as possible with a more heterogeneous sample consisting of students enrolled in diverse study programs. Furthermore,

students' scores on a common exam, gathered about 3 months after the main assessment were added as a longitudinal criterion that was identical for all students. These additions allowed us to investigate the reliability and generalizability of the findings reported in Study 1.

Based on the results found in Study 1, we expected to find CPS to significantly predict both GPA and subjective indicators of university success (Hypothesis 3).

We furthermore expected to find incremental validity of CPS over and above intelligence in predicting university students' GPA and subjective university success. However, the larger variation in university programs should be associated with a larger variation in intelligence. This should increase the validity of intelligence and in turn limit the validity of CPS after controlling for intelligence (Hypothesis 4). This effect should be particularly strong for GPA and less strong for subjective university success (Robbins et al. 2006).

The high heterogeneity in our sample might result in substantial differences in the students' average GPA as well as the frame of reference for the subjective evaluations of their university success. To have an additional common indicator of university success, students' scores on a course final exam taken by all students participating in Study 2 were included. This additional indicator of university success was gathered several weeks after the main assessment thus allowing for a longitudinal prediction of the students' performance. We expected CPS to predict these final exam scores as well (Hypothesis 5).

This effect should not be solely due to intelligence either, and we expected CPS to show incremental validity in the prediction of students' exam scores over and above intelligence (Hypothesis 6). Similar to GPA, the validity of CPS should be reduced by controlling for intelligence.

Method

Participants Data from $N = 216$ students (71% women; age: $M = 23.8$ years; $SD = 5.5$ years) in an obligatory introductory lecture on educational assessment for sophomore teacher-education students at the same German university as in Study 1 were used in Study 2. Most students were enrolled in multiple study programs because teachers in Germany are supposed to teach at least two different school subjects (in addition to the aspects of their study program that are identical for all students striving for teaching degrees, like educational assessment).

Procedure The assessment of demographics and intelligence took place during the first part of the first lecture of the semester. During the following week, the students worked on the CPS tasks online. The final exam covering the entire course took place at the end of the semester. Here, students could choose between two exam dates: either directly at the end of the lecture period (8 weeks after the main assessment) or 2 months later.

Measures The measures used in Study 1 were also administered in Study 2. CPS was assessed with 10 MicroDYN tasks (Greiff et al. 2012). Intelligence was assessed with the IST-Screening (Liepmann et al. 2012). Self-reported university GPAs at the time of the study were reverse-coded so that the best possible passing score was 4 (*very good*) and the worst score was 1 (*sufficient*). The 5-item subjective university success scale ($\alpha = .79$) was completed only by those students who had registered for the second final exam ($n = 181$), and it was filled out immediately prior to beginning the final exam. Performance on the course final exam was used as additional indicator of students' university success. Both exams consisted of 136 multiple-choice items assessing students' competencies and knowledge of

educational and psychological assessment; we used those 58 items that were identical for both exam dates.

Statistical analysis In line with Study 1, the predictions of GPA and subjective university success were calculated in the same models. As students' were enrolled in multiple study programs, we could not account for the within-cluster variance though. The predictions of exam scores were calculated in separate models.

Results

Descriptive statistics and measurement models Table 3 shows the descriptive statistics and intercorrelations for all variables included in Study 2.

Measurement models were established for all latent variables in the same way as in Study 1. All measurement models fit very well to the data thus allowing for estimations of the structural models (Table 4).

Structural models To test Hypothesis 3, CPS was defined to predict both the latent subjective university success factor and manifest GPA scores. This model represented the data very well (Table 4). In accordance with Hypothesis 3, CPS predicted both subjective university success ($\beta = .27$; $p = .038$; $R^2 = .07$) and GPA ($\beta = .15$; $p < .001$; $R^2 = .02$). The correlation between subjective university success and GPA was $r = .62$ ($p < .001$).

To investigate the incremental validity of CPS (Hypothesis 4), CPS was again residualized for intelligence, which explained 70% ($\beta = .83$; $p < .001$) of the variance in CPS. The remaining residual of CPS as well as intelligence itself were then specified to predict both subjective university success and GPA. In support of Hypothesis 4, the residual of CPS remained a significant predictor of both subjective university success ($\beta = .17$; $p < .001$) and GPA ($\beta = .08$; $p < .001$). CPS predicted subjective success significantly stronger than GPA after intelligence was controlled for ($\chi^2 = 3.51$; $df = 1$; $p = .003$). In this model, intelligence itself significantly predicted both subjective university success ($\beta = .41$; $p < .001$) and GPA ($\beta = .18$; $p < .001$). Combined, intelligence and CPS explained 4% of the variance in GPA and 19% of the variance in subjective university success. The correlation between subjective university success and GPA was $r = .63$ ($p < .001$). This model represented the data very well (Table 4).

Table 3 Descriptive statistics and correlations for Study 2

Variables	Mean (SD)	GPA	SUS	Exam scores	CPS acquisition	CPS application	Intelligence
GPA	2.18 (0.47)	–					
SUS	3.51 (0.65)	.52**	–				
Exam scores	44.75 (6.41)	.30**	.23**	–			
CPS acquisition	0.40 (0.27)	.04	.23**	.11	–		
CPS application	0.41 (0.24)	.00	.26**	.09	.71**	–	
Intelligence	48.02 (5.42)	.15*	.46**	.10	.46**	.43**	–

CPS complex problem solving, GPA grade point average, SUS subjective university success

* $p < .05$, ** $p < .01$

Table 4 Model fit indices for Study 2

Model	χ^2	df	<i>p</i>	CFI	TLI	RMSEA
Measurement models						
SUS ($\lambda = .45-.90$)	3.05	5	.27	1.00	1.00	.00
Intelligence ($\lambda = .54-.65$)	0	0	–	1.00	1.00	.00
CPS ($\lambda = .83-.94$)	220.78	170	< .01	.99	.99	.04
Structural models						
CPS predicting GPA and SUS	375.47	298	< .01	.96	.96	.04
CPS predicting GPA and SUS controlling for intelligence	480.72	375	< .01	.97	.97	.03
CPS predicting exam scores	237.71	189	< .01	.99	.99	.03
CPS predicting exam scores controlling for intelligence	290.78	250	< .01	.99	.99	.02

df degrees of freedom, *CFI* comparative fit index, *TLI* Tucker-Lewis index, *RMSEA* root mean square error of approximation, *SUS* subjective university success, *CPS* complex problem solving, *GPA* grade point average

Hypothesis 5 stated that CPS predicted students' exam scores. To test this hypothesis, a latent CPS factor indicated by knowledge acquisition and knowledge application was defined to predict manifest exam scores. CPS predicted exam scores significantly ($\beta = .13$; $p = .038$; $R^2 = .02$). The fit for this model was very good (Table 4).

To control the effect of CPS in the prediction of students' exam scores for intelligence (Hypothesis 6), CPS was residualized for intelligence. In this model, intelligence explained 56% ($\beta = .75$; $p < .01$) of the variance in CPS. Contrary to Hypothesis 5, the remaining residual of CPS no longer significantly predicted students' exam scores ($\beta = .06$; $p = .295$). Intelligence, on the other hand, predicted students' exam scores significantly ($\beta = .15$; $p = .043$). Combined, intelligence and CPS explained 3% of variance in students' exam scores. This model represented the data very well (Table 4).

Discussion of Study 2

The results of Study 2 confirm the role of CPS in the prediction of students' university success. Nonetheless, comparing the results of Study 1 and Study 2 reveals several differences regarding the interplay of CPS and intelligence as predictors of grades and subjective university success. CPS significantly predicted GPA and exam grades as well as subjective university success. However, controlling for intelligence resulted in a substantial drop in the relative importance of CPS in the prediction of all three indicators of university success. After controlling for intelligence, CPS remained a statistically significant predictor of GPA and subjective university success but with considerably lower beta weights than in Study 1. CPS no longer significantly predicted students' exam scores after intelligence was controlled for. This may be due to the exam scores representing a single measure as opposed to the highly aggregated GPA.

Taken together, Study 2 replicated most of the findings of Study 1, however, with substantially smaller effect sizes for CPS and substantially larger effect sizes for intelligence. This supports our interpretation of CPS as a relevant predictor of university success that provides additional information over and above intelligence. The magnitude of this incremental validity may depend on the student population of interest. Future research might investigate student populations of majors with an expressed focus on problem solving such as medicine or engineering.

General discussion

Complex problem-solving skills appear to be relevant for the academic achievement of university students. Both studies presented here corroborate the role of this twenty-first century skill in (tertiary) education with substantial validity. However, the importance of CPS as an additional source of information on students' cognitive ability seems to increase with the selectiveness of university programs (Jensen 1998). This suggests that individual differences in CPS may be helpful in explaining why highly intelligent students still differ in their academic success. Beyond being smart enough to cope with the academic demands of university, students need to learn to extract relevant information, test hypotheses, and control a dynamically changing environment of interrelated variables (Funke 2001) to succeed in their studies at the university level. Regarding Bloom's (1956) taxonomy of educational objectives of the cognitive domain, these skills correspond to higher-order thinking skills and thus the highest aims of education (e.g., evaluation, application).

Limitations

Some limitations need to be considered in the interpretation of the data. The (mostly) cross-sectional design of both studies hampers the causal interpretations of the data. Specifically, the correlation between CPS and indicators of university success may represent an increase in CPS as an outcome of university studies rather than individual differences in CPS causing different levels of university success. However, this limitation holds substantially less for the exam scores (Hypothesis 5), which were assessed a considerable amount of time after CPS. In line with the interpretation of CPS as a predictor rather than an outcome of university success, CPS predicted these exam scores as well. Notably, the validity coefficients found for intelligence in our studies correspond to those reported in longitudinal studies (Hell et al. 2007).

In addition, the choice of operationalization for both CPS and intelligence may have influenced the results. In fact, recent meta-analytic findings have shown that the correlation between CPS and intelligence depends on their operationalization (Stadler et al. 2015b). Measures of CPS and intelligence (such as those used in the current studies) showed the strongest correlations. The incremental validity of CPS may therefore have been stronger using different measures of CPS and intelligence (cf. Kretzschmar et al. 2016; Lotz et al. 2016). Future research may investigate whether other operationalizations of CPS and intelligence lead to a stronger incremental validity of CPS over and above intelligence in the prediction of university success.

Finally, our studies focused exclusively on cognitive predictors of university success. As noted above, various noncognitive predictors of university success have been established as well (Richardson et al. 2012). Little is known about the relation between CPS and noncognitive constructs. Greiff and Neubert (2014) report weak (albeit partly significant) relations between CPS and personality traits. Thus, it stands to reason that future research on the role of CPS in students' university success needs to include noncognitive factors as well to obtain a more complete picture on the various factors influencing students' success in their studies at the university.

Implications

The major finding of both studies presented here is that individual differences in CPS are related to student's university success and that this difference cannot be reduced to individual

differences in intelligence. Next to providing a solid ground for future research on the role of CPS in university success, this may bring practical implications for both selection and teaching in higher education. CPS tasks may represent a valuable addition to other instruments used in university selection. Besides their validity in predicting relevant outcomes, CPS tasks have been shown to be highly accepted by participants (Sonnleitner et al. 2012), whereas intelligence measures suffer from low acceptance in university selection (Hell and Schuler 2005). Minimizing the complexity of university studies by providing guidance and structure on the other hand, may reduce the demands of university programs and allow students to perform better. Finally, we hope that this study instigates attempts on improving students' CPS skills. Helping them to handle the complexity of university studies may thus increase students' success, improve their satisfaction with their education, and ultimately limit the likelihood of a preventable drop-out.

Funding information This research was funded by grants of the Fonds National de la Recherche Luxembourg (ATTRACT “ASKI21”; AFR “COPUS”).

The contribution of the third author was supported by the German funds “Bund-Länder-Programm für bessere Studienbedingungen und mehr Qualität in der Lehre (‘Qualitätspakt Lehre’)” [the joint program of the federal government and states for better study conditions and the quality of teaching in higher education (“the Teaching Quality Pact”)] at Saarland University (funding code 01PL11012). The concept and content of this manuscript was developed by the authors independently of this funding.

References

- Arnett, J. J. (2000). Emerging adulthood: a theory of development from the late teens through the twenties. *American Psychologist*, *55*, 469–480. <https://doi.org/10.1037//0003-066X.55.5.469>.
- Bacon, D. R., & Bean, B. (2006). GPA in research studies: an invaluable but neglected opportunity. *Journal of Marketing Education*, *28*, 35–42. <https://doi.org/10.1177/0273475305284638>.
- Beauducel, A., & Wittmann, W. W. (2005). Simulation study on fit indexes in CFA based on data with slightly distorted simple structure. *Structural Equation Modeling*, *12*, 41–75. https://doi.org/10.1207/s15328007sem1201_3.
- Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives: Vol. 1. Cognitive domain*. New York: McKay.
- Duckworth, A. L., Weir, D., Tsukayama, E., & Kwok, D. (2012). Who does well in life? Conscientious adults excel in both objective and subjective success. *Frontiers in Psychology*, *3*, 1–8. <https://doi.org/10.3389/fpsyg.2012.00356>.
- Ferrett, S. (2000). *Peak performance: Success in college and beyond*. New York: Glencoe/McGraw-Hill.
- Formazin, M., Schroeders, U., Koeller, O., Wilhelm, O., & Westmeyer, H. (2011). Studierendenauswahl im Fach Psychologie. Testentwicklung und Validitätsbefunde student selection for psychology. Test development and predictive validity. *Psychologische Rundschau*, *62*, 221–236.
- Frensch, P. A., & Funke, J. (Eds.). (1995). *Complex problem solving: the European perspective*. Hillsdale: Erlbaum.
- Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking & Reasoning*, *7*, 69–89. <https://doi.org/10.1080/13546780042000046>.
- Funke, J., & Frensch, P. A. (2007). Complex problem solving: the European perspective—10 years after. In D. H. Jonassen (Ed.), *Learning to solve complex scientific problems* (pp. 25–47). New York: Lawrence Erlbaum.
- Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2003). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences*, *14*, 47–64. <https://doi.org/10.1016/j.lindif.2003.08.002>.
- Gattiker, U. E., & Larwood, L. (1986). Subjective career success: a study of managers and support personnel. *Journal of Business and Psychology*, *1*, 78–94.
- Greiff, S., & Neubert, J. C. (2014). On the relation of complex problem solving, personality, fluid intelligence, and academic achievement. *Learning and Individual Differences*, *36*, 37–48. <https://doi.org/10.1016/j.lindif.2014.08.003>.
- Greiff, S., Wüstenberg, S., & Funke, J. (2012). Dynamic problem solving: a new assessment perspective. *Applied Psychological Measurement*, *36*, 189–213. <https://doi.org/10.1177/0146621612439620>.

- Greiff, S., Wüstenberg, S., Molnár, G., Fischer, A., Funke, J., & Csapó, B. (2013). Complex problem solving in educational contexts—Something beyond g: concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology*, *105*, 364–379. <https://doi.org/10.1037/a0031856>.
- Greiff, S., Stadler, M., Sonnleitner, P., Wolff, C., & Martin, R. (2015a). Sometimes less is more: comparing the validity of complex problem solving measures. *Intelligence*, *50*, 100–113. <https://doi.org/10.1016/j.intell.2015.02.007>.
- Greiff, S., Fischer, A., Stadler, M., & Wüstenberg, S. (2015b). Assessing complex problem-solving skills with multiple complex systems. *Thinking & Reasoning*, *21*, 356–382. <https://doi.org/10.1080/13546783.2014.989263>.
- Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: a longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology*, *94*, 562–575. <https://doi.org/10.1037/0022-0663.94.3.562>.
- Hell, B., & Schuler, H. (2005). Verfahren der Studierendenauswahl aus Sicht der Bewerber. [Methods of university student selection from the applicants' view]. *Empirische Pädagogik*, *19*(4), 361–376.
- Hell, B., Trapmann, S., & Schuler, H. (2007). Eine Metaanalyse der Validität von fachspezifischen Studierfähigkeitstests im deutschsprachigen Raum [A meta-analytic investigation of subject-specific admission tests in the German-speaking countries]. *Empirische Pädagogik*, *21*, 251–270.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. <https://doi.org/10.1080/10705519909540118>.
- Huang, F. L. (2016). Alternatives to multilevel modeling for the analysis of clustered data. *The Journal of Experimental Education*, *84*, 175–196. <https://doi.org/10.1080/00220973.2014.952397>.
- Hughes, E. C. (1959). Men and their work. *American Journal of Sociology*, *65*(1), 115–117.
- Jensen, A. (1998). *The g factor: the science of mental ability*. Westport: Praeger.
- Johnson, V. E. (2003). *Grade inflation: A crisis in college education*. New York: Springer.
- Kretzschmar, A., Neubert, J. C., Wüstenberg, S., & Greiff, S. (2016). Construct validity of complex problem solving: a comprehensive view on different facets of intelligence and school grades. *Intelligence*, *54*, 55–69. <https://doi.org/10.1016/j.intell.2015.11.004>.
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: Implications for graduate student selection and performance. *Psychological Bulletin*, *127*, 162–181. <https://doi.org/10.1037/0022-3514.86.1.148>.
- Liepmann, D., Beauducel, A., Brocke, B., & Nettelstroth, W. (2012). *Intelligenz-Struktur-Test—screening—IST-screening [intelligence structure test—screening]*. Göttingen: Hogrefe.
- Lotz, C., Sparfeldt, J. R., & Greiff, S. (2016). Complex problem solving in educational contexts—still something beyond a “good g”? *Intelligence*, *59*, 127–138. <https://doi.org/10.1016/j.intell.2016.09.001>.
- Muthén, L. K., & Muthén, B. O. (1998–2015). Mplus user's guide. Seventh Edition. Los Angeles: Muthén & Muthén.
- OECD. (2014). *Education at a glance 2014: OECD indicators*. Paris: OECD Publishing. <https://doi.org/10.1787/eag-2014-en>.
- Parker, J. D., Summerfeldt, L. J., Hogan, M. J., & Majeski, S. A. (2004). Emotional intelligence and academic success: examining the transition from high school to university. *Personality and Individual Differences*, *36*, 163–172. [https://doi.org/10.1016/S0191-8869\(03\)00076-X](https://doi.org/10.1016/S0191-8869(03)00076-X).
- Pittman, L. D., & Richmond, A. (2008). University belonging, friendship quality, and psychological adjustment during the transition to college. *The Journal of Experimental Education*, *76*, 343–362. <https://doi.org/10.3200/JEXE.76.4.343-362>.
- Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychological Bulletin*, *138*, 353–387. <https://doi.org/10.1037/a0026838>.
- Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*, 261–288. <https://doi.org/10.1037/0033-2909.130.2.261>.
- Robbins, S. B., Allen, J., Casillas, A., Peterson, C. H., & Le, H. (2006). Unraveling the differential effects of motivational and skills, social, and self-management measures from traditional predictors of college outcomes. *Journal of Educational Psychology*, *98*, 598–616. <https://doi.org/10.1037/0022-0663.98.3.598>.
- Sonnleitner, P., Brunner, M., Greiff, S., Funke, J., Keller, U., Martin, R., et al. (2012). The genetics lab. Acceptance and psychometric characteristics of a computer-based microworld to assess complex problem solving. *Psychological Test and Assessment Modeling*, *54*, 54–72.
- Stadler, M. J., Becker, N., Greiff, S., & Spinath, F. M. (2015a). The complex route to success: complex problem-solving skills in the prediction of university success. *Higher Education Research & Development*, *35*, 1–15. <https://doi.org/10.1080/07294360.2015.1087387>.
- Stadler, M., Becker, N., Gödker, M., Leutner, D., & Greiff, S. (2015b). Complex problem solving and intelligence: a meta-analysis. *Intelligence*, *53*, 92–101. <https://doi.org/10.1016/j.intell.2015.09.005>.

- Sternberg, R. J., & Berg, C. A. (1986). Quantitative integration. Definition of intelligence: a comparison of the 1921 and 1986 symposia. In R. J. Sternberg & D. K. Detterman (Eds.), *What is intelligence?* (pp. 155–162). Norwood: Ablex.
- Strenze, T. (2007). Intelligence and socioeconomic success: a meta-analytic review of longitudinal research. *Intelligence*, 35, 401–426. <https://doi.org/10.1016/j.intell.2006.09.004>.
- Tavernier, R., & Willoughby, T. (2014). Bidirectional associations between sleep (quality and duration) and psychosocial functioning across the university years. *Developmental Psychology*, 50, 674–682. <https://doi.org/10.1037/a0034258>.
- Terenzini, P. T., Rendon, L. I., Upcraft, M. L., Millar, S. B., Allison, K. W., Gregg, P. L., & Jalomo, R. (1994). The transition to college: diverse students, diverse stories. *Research in Higher Education*, 35, 57–73. <https://doi.org/10.1007/BF02496662>.
- Wagener, D. (2001). *Psychologische Diagnostik mit komplexen Szenarios [Psychological diagnostics using complex scenarios]*. Lengerich: Pabst.
- Wirth, J., & Klieme, E. (2003). Computer-based assessment of problem solving competence. *Assessment in Education: Principles, Policy, & Practice*, 10, 329–345. <https://doi.org/10.1080/0969594032000148172>.
- Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving—more than reasoning? *Intelligence*, 40, 1–14. <https://doi.org/10.1016/j.intell.2011.11.003>.