Chracterizing successful participants in ScienceOlympiads – Evidence from the PhysicsOlympiad

Peter Wulff[a], Stefan Petersen[b], Melanie Keller[b], Tim Höffler[b], and Knut Neumann[b]

[a]Physics Educational Research Group, University of Potsdam, Karl-Liebknecht-Straße 24/25, 14476 Potsdam-Golm
[b]Leibniz Institute for Science and Mathematics Education, Olshausenstrasse 62, 24118 Kiel, Germany

ARTICLE HISTORY
Compiled March 5, 2019

ABSTRACT
Given the need for high-achieving students to engage in science, technology, engineering, and math (STEM), this study seeks to characterize successful students in the Physics Olympiad as a means to enable future educational efforts to be more in congruence with the characteristics of the students. On the basis of the expectancy-value model of achievement motivation and research in expertise, $N = 141$ students were tracked in their engagement with the Physics Olympiad and administered appropriate motivational and cognitive constructs. The dependent outcome variable was the success the students had in their participation in the Physics Olympiad. Results indicate that successful students can be characterized through high skills in physics problem solving and positive motivational attributes, namely a high expectation to be successful in the Physics Olympiad. These results pave the path to more transparency for what comprises expertise in domains like physics such that designated educational efforts can motivate and foster more students.

KEYWORDS
Problem solving, Physics competitions

1. Introduction

Enrichment programs such as the ScienceOlympiads are means to identify and promote talented students in science, technology, engineering, and math (STEM). These programs proceed in subsequential stages where students solve increasingly complex domain-specific problem and eventually meet and compete with one each other (Petersen & Wulff, 2017). Amongst the most successful students, a national team is chosen, comprising about five students, who compete on an international level against students from more than 80 countries in the Physics Olympiad. Federal government and the STEM community endorse these means as viable instruments to foster talented students (KMK, 2009; Petersen & Wulff, 2017)–and educational researchers in gifted programs motivated the necessity of ScienceOlympiads (and programs alike) as a viable complementary to regular schools that have limited capacities to provide resources for talented students (Reis & Renzulli, 2010). Besides these goals, research is scare on these programs (Ziegler, 2004).

CONTACT Peter Wulff. Email: peter.wulff@uni-potsdam.de

The available studies document that successful candidates report a positive impact on their future job aspirations in STEM through programs such as the Science-Olympiads (Feng, Campbell, & Verna, 2001; Oswald, Hanisch, & Hager, 2004; Subotnik, Duschl, & Selmon, 1993). Further research suggests that these programs can have effects for training skills related to cognitive abilities and related to beliefs such as developing interest and motivation of students towards STEM (Aljughaiman & Ayoub, 2012; Marsh, Chessor, Craven, & Roche, 1995; Oswald et al., 2004; Wai, Lubinski, Benbow, & Steiger, 2010). Due to the broad motivation of enrichment programs (foster gifted students) and the self-selective mechanisms for participation, other studies sought to characterize participants in these programs in order to advance an understanding for characteristics of successful participants. Urhahne, Ho, Parchmann, and Nick (2012) found for the ChemistryOlympiad that previous participation was the best predictor for success in this competition and also expectancy of success distinguished successful participants from less successful participants (similar findings in: Stang, Urhahne, Nick, & Parchmann, 2014). Taken together, the above studies suggest that successful students in programs such as the ScienceOlympiads show advantageous dispositions in cognitive variables such as general cognitive abilities, and that more successful students display advantageous beliefs such as a high expectancy of success towards the competition.

However, two problems arise in the context of the above studies. First, even though cognitive abilities appear to be particularly predictive for success in these programs, operationalization of domain specific abilities is insufficient so that it remains unclear to what extent domain-specific cognitive abilities are characteristic of successful participants in these programs. Second, no such analyses have been done for the Physics Olympiad. Physics is often suggested to be particularly heavy in content dependency such that characterizing successful participants in the Physics Olympiad might hinge on an integrative assessment of both cognitive variables and beliefs.

2. Modelling success in ScienceOlympiads

Applied to the context of the Science Olympiads, Urhahne et al. (2012) applied the expectancy-value model of achievement motivation in order to explain variance in success for the participants. The expetcancy-value model outlines two proximal causes for achievement related choices and performance in a situation: expectancy to be successful in a task ("Can I do this?") and the values brought towards performing the relevant tasks ("Do I want to do this?") (Eccles, 1983). The model has been empirically validated in multiple contexts such as occupational choices REF, academic choices in school (Köller, Daniels, Schnabel, & Baumert, 2000) and ScienceOlympiads (Urhahne et al., 2012).

Expectancy of success and values towards the context (e.g., ScienceOlympiad) are also influenced by other multiple variables that commonly in talent research relate to cognitive dispositions (stable and vairable), affective/motivational variables, and external motivational moderators (e.g., Heller, 2002; Ziegler & Stoeger, 2009). Regarding cognitive dispositions, it has been implicated from the inception of talent studies that general cognitive abilities successful from less successful students–talent was even defined through scores in general cognitive abilities (Rost, 2010). Later on, incited by studies in domains such as chess and physics, researchers acknowledged that domain-specific skills such as problem solving are the characteristic features that distinguish successful from less successful students in a domain (**?**). Today, it is clear that successful students in a domain invest enormous amounts of deliberate practice to master a domain (Simon, 1983). Consequently, domain-specific abilities are essential for characterizing successful students in programs such as the ScienceOlympiads.

Affective and motivational variables can be characterized to be beliefs about oneself that potentially impact goal-directed behavior in that domain (Rheinberg & Vollmeyer, 2012). There is contentious debate, what essential affective and motivational aspects direct human behavior, but competence, belongingness, and interest are conceptualized in most models (Deci & Ryan, 2000; Hazari, Sonnert, Sadler, & Shanahan, 2010). Bandura (1997) presented a broad program outlining the importance of self-efficacy beliefs (closely linked to competence) and educational outcomes such as positive attitudes towards performance. Regarding belongingness, Baumeister and Leary (1995) muster evidence that humans perform better when they feel accepted and related to people in the respective community. Good, Rattan, and Dweck (2012) link the sense of belonging for students to academic choices in mathematics. Finally, domain interest is linked to positive learning outcomes and academic choices in the domain (Hazari et al., 2010; Krapp, 2002). External motivational moderators can be conceived of as social support by meaningful others. For example, the attitudes of peers, teachers, and parents are an important facilitator for achievement in an academic competition (e.g., Urhahne et al., 2012).

While for the ChemistryOlympiad and the BiologyOlympiad first results with regards to characterization of successful participants are available (Stang et al., 2014; Urhahne et al., 2012), such results are missing for the PhysicsOlympiad. Furthermore, instruments for measuring domain-specific abilities are missing such that the characterization of successful participants to date remains incomplete. Consequently, two research questions are addressed in this study:

(1) To what extent can cognitive variables (general cognitive abilities and domain-specific abilities) characterize successful students in the PhysicsOlympiad?
(2) To what extent do affective/motivational variables additionally explain success in the PhysicsOlympiad?

3. Design

Characterizing successful students in the PhysicsOlympiad suggests a design where students' performance in the competition is registered and where (in the optimal case all) participating students respond to the above outlined constructs. This is facilitated through two measurement times where students respond to items on an online platform. As soon as participants registered in the PhysicsOlympiad registration platform (each year approx. 1000 students do so (Petersen & Wulff, 2017)), these students received an invitation to participate in an online survey that accompanies them throughout their participation in the competition and where the students could win prices through participation. Participation was entirely voluntary, anonymous, and independent from the competition participation. This information was given to the students prior to their participation. Throughout the students' participation advancing rounds of the online questionnaire were unlocked and open for participation.

At some point during their participation in stage 1 of the PhysicsOlympiad (where students do homework problems and hand them in to their physics teacher) the participants received an email that the first online questionnaire was open for participation. Students who did not participate received reminder emails until at some point the first

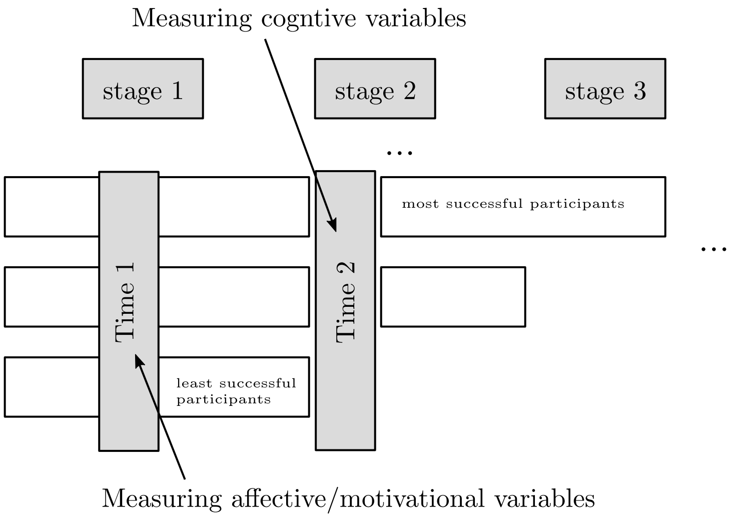


Figure 1. Design for measurements of different variables throughout the PhysicsOlympiad.

questionnaire was closed. In the parallel competition the students received a notification whether they advanced through stage 2 in the competition. After this notification the second online questionnaire was unlocked for participation. This procedure repeated until after stage 4. The relevant design for the current study is displayed in Figure 1.

4. Instruments

Instruments are employed with regards to cognitive, affective/motivational, and external motivational variables.

4.1. Cognitive variables

Cognitive variables that potentially predict success in programs such as the ScienceOlympiads are general cognitive abilities. A simplified version of Carroll's hierarchical theory of cognitive abilities is the radex model where quantitative/numerical, spatial/pictorial, and verbal/linguistic abilities are differentiated (Snow, Corno, & Jackson, 1997). Students that excel in STEM show generally high cognitive abilities. For example, Wai, Lubinski, and Benbow (2009) find that less than 10 percent of those holding a STEM-PhD were below the top quartile in cognitive ability (comprising spatial visualization in 2D and 3D, mechanical reasoning, and abstract reasoning) during adolescence, and their data supports the conclusion that the importance of spatial ability for STEM increases with more advanced degrees (bachelor, master, PhD). Graduates in physical sciences such as engineering range among the top performers in cognitive abilities (see also: Shea, Lubinski, & Benbow, 2001). "STEM disciplines place a premium on nonverbal ideation indicative of quantitative and spatial reasoning" (Lubinski, 2010). Consequently, this study focuses on quantitative and numerical abilities, because of its central importance for STEM (Wai et al., 2009). A sample item from the test for general cognitive abilities in numerical scale asks: "What quantity is bigger: $\sqrt{100}$ or 50?" 34 items were used from the quantitative and numerical scales from the test for general cognitive abilities (Heller & Perleth, 2007). Only items that were applicable for grades 9 to 12 were retrieved to build a sum score for each participant. This gives an approximation of general cognitive abilities with regards to a norm, rather than an age adaptive measure as is usually done in intelligence testing.

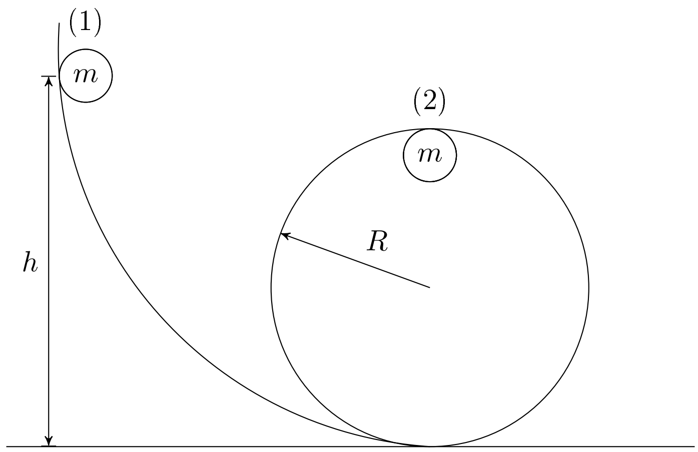
Besides general cognitive abilities, indicative of success in the PhysicsOlympiad is likely to be physics problem solving abilities (Maloney, 2011). Physics problem solving is laregly based on prior knowledge (Bhaskar & Simon, 1977), where commonly factual knowledge (declarative), strategic (procedural) knowledge, and conditional knowledge are necessary (Anderson, Reder, & Lebiere, 1996), which are chunked in problem schemata (Reinhold, Lind, & Friege, 1999). Knowledge and problem schemata are applied in different phases of problem representation, strategy selecton, solution, and verification of solution (Mayer, 2013; Polya, 1945). Experts carefully conduct a qualitative problem analysis prior to solving a problem based on the fundamental principles involved (Larkin, McDermott, Simon, & Simon, 1980; **?**). Expert knowledge is furthermore applicable in new situations (conditionalized) (Glaser, 1992; Simon, 1980), because experts categorize problems by basic concepts (deep structure), whereas novices categorize problems by surface features such as "is a pulley involved?" (**??**). Novices have problems to recognize and reflect upon the assumptions that predicate the solution to a problem. Reif and Allen (1992) report that novice students showed deficient applicability conditions for concepts and thus the use of the concepts was flawed.

Physics problem solving ability, here, entails the ability to apply explicit, structured knowledge in various problem situations to achieve a desired goal state. Leonard, Dufresne, and Mestre (1996) propose a measure for physics problem solving ability that assesses the strategies that students use to solve problems. Strategies, as the authors use the term, entail the concept to solve the problem, a justification for why the concepts can be applied (i.e., conditional knowledge), and a procedure by which the concept is applied. The authors emphasize the role of conceptual knowledge that is important for physics problem solving and the idea that problem solving strategies shall be established before applying mathematics. On the basis of the aforementioned definition of physics problem solving ability, a similar rationale as in citeLeonard.1996 was adopted in this study in order to measure physics problem solving ability. Students were asked to write down the physics ideas that they would use to solve the well-defined physics problem without going through the whole problem solution process (i.e., leaving out problem representation and mathematical routines). Consequently, the items provided the students space to construct their understanding and give context to their thoughts (Hammann & Jördens, 2014). In order to make clear to students what was meant with the prompt, an example was provided to all students that they read through. Afterwards, four physics problems (two mechanics, two electro-magnetism) followed that each had a certain context and well-defined solution plans.

A theory-based coding rubric was utilized for scoring the responses. Docktor et al. (2016) developed a similar coding rubric for students' written solutions for physics problems. Not all of the five categories were adopted for the measure in this study, so that the coding rubric comprised the categories concept, execution, context, and detail. Concept entails information if students provide all relevant concepts that are necessary for solving the problem (physics approach in Docktor et al. (2016)). Execution entails information about the comprehensiveness of the response (logical progression). Context reflects the assumptions and conditions for applying the concepts (specific application of physics) (Fortus, 2009), and, finally, detail detected when students' interspersed their response with an elaboration of the concepts such as through mathematical formulas where the relevant variables became clear. Furthermore, these aspects also relate to the strategy definiton in Leonard et al. (1996). Similar to the rubric by Docktor et al. (2016), all categories were graded with 0 to 2 points. Moreover, concept and context could be graded negative (-1) when a concept was wrong (e.g., when students used magnetic force instead of electric force in a problem). For concept the full score was given when the student response included all concepts that were central to the respective problems. Execution pertained the presentation of the physical ideas. If the sentences were understandable (not only two-word phrases) and logically connected, full credit was given. Full credit for detail was awarded when all relevant variables in the concepts were identified. Full credit for context was given when important assumptions and boundary conditions were identified (e.g., frictionless slide). Note that the categories are not idependent of each other. As a consequence, we aggregated the score for each student to an overall measure which is henceforth called ability of problem conceptualization, because a full credit (i.e., 8 points for each problem) would indicate that the student has all the relevant analyses correct that would count for successful physics problem solving.

Four physics problems were presented to the students that are related to two fundamental physics topics that are encountered throughout physics: mechanics and electromagnetism. Knowledge for gravitational and electro-magnetic force, momentum, and energy conservation were necessary to solve all problems. A sample problem from mechanics reads:

A mass runs through a looping (see picture below). The mass starts from a heights above the highest point of the looping. The movement is frictionless. Determine the minimal necessary heights for the mass, so that the mass can run through the looping without falling down. Assume that the mass does not slip and that it is a point mass *Describe how you would solve this problem and what physics ideas you would use to solve it. Try to write full sentences.*



An expert answer for this open-ended problem looked as follows:

> The centrifugal force at the highest point of the looping needs to be equal with the weight: $(m * v^2)/r = m * g$. In order to figure out how high the starting point needs to be, we calculate the initial speed that he needs in order to pass the looping and with the energy conservation law we determine the height: $v = \text{sqrt}(g * r)$, $mgh = m/2 * v^2 h = 0.5r + 2r$ (height of the looping) $h = 5/2R$.

In this answer, 2 points were given for the concepts (equilibrium of forces and energy conservation). Furthermore the student's answer is coherent and it is easy for the reader to follow (2 points for execution). Details are given through the correct formulas. Missing points appear for the context, because the student does not include a reflection

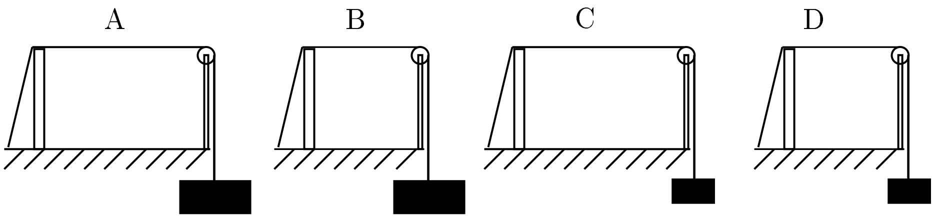


Figure 2. Sample item from general physics problem solving abilities test from Heller and Perleth (2007). Choice E was the neutral choice "Cannot be decided."

of boundary conditions in the response. A novice response, were no credit is awarded, looked like this:

> I would recognize the ascent of the trajectory and the height of the looping. Besides, I think the mass of the object is crucial. At first I would calculate the velocity or acceleration on the initial path and the velocity with which the object enters the looping. With the gravitational acceleration and the initial acceleration you could calculate which acceleration and initial height is necessary that the object does not fall out the looping.

In order to validate the ratings, two raters (first author and physics undergraduate student with extensive experience in math and physics) that were familiarized with the problems coded the student responses. Both raters were trained to score the responses with regards to the rubric. All student responses were scored. When both raters disagreed, the disagreements were discussed and eventually a consensus was reached with regards to the final scoring.

The final agreement as measured through Cohen's kappa was 0.95. These values are satisfatory (Bortz & Döring, 2002). The retest reliability for the strategy knowledge measure was 0.54 ($N = 101$).

Finally, a measure for more general physics problem solving abilities was employed that was utilized in other gifted research as well Heller and Perleth (2007). Keeping the same physics problem solving ability definition as above, this measure is more closely linked to an intuitive understanding of physics systems. Students were given multiple choices as response alternative and had to choose the correct answer. A sample item can be seen in Figure 2 (correct answer is C). Overall, this test comprises 15 items (each multiple choice with each 5 responses) that had to be solved in 6 minutes. Sum scores were retrieved for students.

4.2. Affective/motivational variables

Affective/motivational and external motivational variables are gleaned from the expectancy-value model and further research on motivated behavior. Expectancy of success towards the PhysicsOlympiad was measured with 5 item such as "I think that I will be successful in the PhysicsOlympiad." Expectancy of success is measured on a 4-point Likert scale with anchors "not true" to "exactly true". Internal consistency is high (see Table 1).

The expectancy-value model of achievement motivation predicts also that values towards tasks in the domain are relevant for engagement in the PhysicsOlympiad. However, these effect do only show partially (for utility value) in reality (Urhahne et al., 2012). In this study, an aggregate measure of values (attainment, cost, utility, and interest) was used, because internal consistency across these items is high (see Table

1). Values are measured with 4 items such as "It is personally relevant for me to be good at the PhysicsOlympiad." Values are measured on a 4-point Likert scale with anchors "not true" to "exactly true".

A related construct to expectancy of success in the PhysicsOlympiad is the self-efficacy in physics (i.e., the perception to be able to solve physics problems). Self-efficacy is a predictor for school success (e.g., Britner & Pajares, 2006) and more broadly construed as compared to expectancy of success in the PhysicsOlympiad. Self-efficacy is measured with 4 items such as "I am certain to understand the difficult contents in physics." Response scale is 4-point Likert with anchors "not true" to "exactly true."

Of yet unexplored importance is the sense of belonging to physics. It is documented that sense of belonging (the perceived belonging to a domain) is important for students to build the intention to study STEM (Good et al., 2012), however, it is unclear if sense of belonging can explain further variance between students who engage in a ScienceOlympiad. Sense of belonging scale was adopted from Good et al. (2012). Items were translated into German and 3 well functioning items (as evidenced through psychometric indicators in the study by Good et al. (2012)) from each subscale were retained in the scale. Overall, sense of belonging was measured with 15 items such as "In physics environments I trust that test materials are not bias such that they benefit particular students." Again, a 4-point Likert scale with anchors "not true" to "exactly true" was utilized. Internal consistency of the scale was also high.

As an external motivational variable, social support by peers, teachers, and parents was utilized. Perceived social support was an important predictor for general interest in physics (Hoffmann, 2002). Consequently, social support was included in analyses with 12 items such as "My parents/teacher/peers supported me with regards to my participation in the PhysicsOlympiad." Response scale was a 4-point Likert scale with anchors "not true" to "exactly true".

Success criterium was highest achieved round of the participants in the PhysicsOlympiad (see Figure 1). The number of students in the different stages was as follows: stage 1: 66 (44%), stage 2: 64 (57%), stage 3: 11 (92%), stage 4: 3 (38%). In brackets are the percentage of students in this analysis from the overall sample that took part in the online questionnaire. Particularly in stage 3, a large perentage of students participated also in the problem solving task. Due to the small overall number of students, this could be a mere coincidene.

Table 1. Overview of utilized variables.

Scale	N	#	α	$\bar{r}_{i(t-i)}$	Scale	M	SD
Cognitive variables:							
Phys. probl. sol. ability	144	16	0.87	0.51	0,1,2	7.46	6.78
General phys. prob. sol. ability	133	15	0.69	0.29	0.15	7.05	2.83
General cog. abilities	78	34	0.88	0.4	0,1	23.72	7.13
Affective/motivational variables:							
Self-efficacy physics	144	4	0.85	0.69	1,2,3,4	3.33	0.57
Expectancy physics competition	144	5	0.71	0.47	1,2,3,4	2.6	0.49
Value physics competition	144	4	0.75	0.55	1,2,3,4	2.85	0.64
Sense of belonging	144	15	0.81	0.42	1,2,3,4	3.32	0.35
Social support in physics	144	12	0.77	0.41	1,2,3,4	2.56	0.5
Success criterium:							
Highest round	144	1			1,2,3,4	1.66	0.71

[a] N ... Number of student responses with at least one valid value, α ... internal consistency and $\bar{r}_{i(t-i)}$ ... mean discriminatory power of items.

5. Sample

The sample comprises students who participated in the online questionnaire and also submitted answers to the physics problem solving abilities items. 144 (m=101, f=43) students responded to these items. Mean (SD) age for the participants was 16.2 (1.2). In order to rule out potential selection biases with regards to the entire population of participants in the online questionnaire, background variables (gender and age) and affective/motivational variables were used to predict group membership through a logistic regression with group as outcome variable. Significant coefficients for a predictor variable would suggest that a certain predictor distinguishes the two groups. No effect became significant (Self-efficacy physics: $\beta = -0.08, SE(\beta) = 0.12, z = -0.63, p = .53$; Expectancy physics competition: $\beta = 0.09, SE(\beta) = 0.13, z = 0.67, p = .50$; Value physics competition: $\beta = 0.03, SE(\beta) = 0.12, z = 0.3, p = .76$; Sense of belonging: $\beta = -0.02, SE(\beta) = 0.12, z = -0.15, p = .88$; Social support in physics: $\beta = 0.01, SE(\beta) = 0.11, z = 0.1, p = .92$; Qualified round: : $\beta = -0.21, SE(\beta) = 0.15, z = -1.36, p = .17$; Gender: $\beta = 0.11, SE(\beta) = 0.25, z = 0.46, p = .65$; Age: $\beta = 0.01, SE(\beta) = 0.1, z = 0.12, p = .90$).

6. Results

Since the affective/motivational variables are self-reported measures and the cognitive variables are expected to be positively related (see positive manifold above), it is necessary to investigate the correlations among the variables first in order to verify that no multi-collinearity is present. Intercorrelations have values in the range between $-0.18 - 0.42$. Mean (SD) correlation is 0.13 (0.15). However, most values are lower such that no multicollinearity is present. One pattern in the correlation matrix is that the cognitive variables correlate significantly amongst each other and the affective/motivational variables correlate significantly amongst each other in positive directions. Social support shows almost no significant correlation with another variable. Interestingly, self-efficacy in physics is not correlated with cognitive variables, which is quiet counter-intuitive given the literature on self-efficacy (see: Bandura, 1997). Expectancy of success in the competition, however, is significantly correlated with the cognitive variables, possibly because it is more context-specific such that students can better judge their actual ability.

In order to characterize successful participants (RQ 1), generalized linear models that account for the ordinal metric of the dependent variable are utilized. Cognitive and affective/motivational variables are included as predictors for highest qualified round (dependent variable). Achievement in round 1 in the physics Olympiad was omitted as a cognitive variable because of the high correlation with the other cognitive measures that were of more interest to us. Proportional odds models were used in order to predict success in the Physics Olympiad in a hierarchical regression procedure (Field, 2012). First, cognitive variables were included as predictors in the model because these predictors can be considered most important for characterizing successful students in the Physics Olympiad based on prior research (e.g., Urhahne et al., 2012). In a second step, cognitive variables and affective/motivational variables were jointly entered in the model. Ordinal regression models furthermore assume parallel slopes (Bürkner & Vuorre, 2018), $\chi^2 = 2.53, p = .47$.

The interesting parameters in the model are the effects (not the intercepts). The effects can be interpreted in accordance with the regression parameters in linear regression (Bürkner & Vuorre, 2018). Table 2 shows the results from the analysis. As can be expected, phys. problem solving ability and generic physics problem solving are significantly positively related with highest qualified round of the participants. The relationship holds when entering the affective/motivational variables to the model. When all variables are included in the model, expectancy of success in the competition shows a significant positive relationship with highest qualified round as well. No other effects become significant and thus characterize successful students in this study. Phys. problem solving ability has the highest odds ratio, and thus a change in phys. problem solving ability of one standard deviation has the highest contribution to the likelihood of a student to reach a higher round.

Table 2. Proportional odds model for predicting highest qualified round in the physics Olympiad with various predictors.

	β	$SE(\beta)$	z	p	OR
Phys. probl. sol. ability	0.71	0.2	3.62	$< .001$	2.04
General phys. prob. sol. ability	0.37	0.18	2.05	$< .05$	1.45
Gen. cog. abilities	0.22	0.19	1.15	.25	1.24
R^2_{adj}	0.23				
Phys. probl. sol. ability	0.7	0.22	3.24	$< .001$	2.01
General phys. prob. sol. ability	0.37	0.19	1.93	$< .05$	1.45
Gen. cog. abilities	0.14	0.19	0.73	.47	1.15
Self-efficacy physics	0.08	0.21	0.36	.72	1.08
Expectancy physics competition	0.57	0.22	2.64	$< .01$	1.77
Value physics competition	0.12	0.19	0.6	.55	1.12
Sense of belonging	-0.33	0.21	-1.59	.11	0.72
Social support in physics	0	0.19	-0.01	1.00	1
Age	-0.11	0.2	-0.52	.60	0.9
R^2_{adj}	0.29				

7. Discussion

In order to better understand how enrichments programs such as the ScienceOlympiads identify and promote their participants, sound empirical evaluation of these programs is inevitable. A first step for evaluation was to characterize successful participants in these programs so potential selective mechanisms of these programs can be uncovered, e.g., if only students from high-status parental background (for whatever reason) make it to the top this might raise threats to the aims and goals of these programs to identify and promote the most talented students. The ChemistryOlympiad and the BiologyOlympiad have been analyzed with this focus. However, an equivalent evaluation for the PhysicsOlympiad is pending and particularly competence measures such as domain-specific problem solving abilities for high-achieving students in the olympiads could not be adopted in the former evaluation studies. This study shought to characterize successful students in the Physics Olympiad with a focus on domain-specific problem solving abilities. The theoretical backbone was the expectancy-value model that has been utilized to evaluate the Chemistry- and BiologyOlympiad as well. Cognitive and affective/motivational variables were gleaned from the expectancy-value model of achievement motivation and from expertise research. Cognitive variables include measures that were not all readily available from previous research. In particular, expertise research suggests that experts outperform novices in problem solving. Consequently, a physics problem measure was developed on the basis of relevant theory and prior measures for problem solving and general cognitive abilities were utilized. The developed measure for physics problem conceptualization ability shows satisfying item characteristics with regards to internal consistency as measured through Cronbach's α, discriminatory power and interrater reliability. Furthermore, discriminatory validity as measured through correlation with similar constructs (?) was satisfactory because the intercorrelations can be considered small to medium size. Affective/motivational variables include, amongst others, expectancy of success, values towards the competition, sense of belonging, and social support for physics engagement.

As a means to answer the research questions and thus characterize successful students in the Physics Olympiad, an ordinal regression was used with a hierarchical inclusion procedure of sets (cognitive and affective/motivational) predictor variables (?). First, the cognitive variables were included first in the model because prior research indicated that they were more important for success (e.g., Urhahne et al., 2012). In a second step, the cognitive and the affective/motivational variables were included together in the model. Physics problem solving ability was found to be significantly related with success in the Physics Olympiad, as assessed through highest qualified round. Besides physics problem solving ability, also what was called general physics problem solving ability proved to be significantly related with success in the Physics Olympiad. Both, physics problem solving ability and general physics problem solving ability remained significantly related with success when the affective/motivational variables were added into the model. From the set of affective/motivational variables, expectancy of success was significantly related to success in the Physics Olympiad. This is consistent with the findings from the Chemistry Olympiad (Urhahne et al., 2012), and with more general findings related the expectancy of success (?). No other variables proved to be significant predictors for success.

Limitations in interpretability and generalizability arise with respect to potential selection bias with regards to the general olympian population. It should be noted that only a fraction of the students that participated in the Physics Olympiad also took part in the online questionnaire. Subjects in this study are the same age as the overall population of olympians M(study) $= 11.18$, $SD = 0.98$; M(population) $= 11.17$, $SD = 1.24$; $t(204.74) = 0.02$, $p = .98$, $r = 0$, however, subjects in this study perform better with regards to participants in general population M(study) $= 27.12$, $SD = 10.38$; M(population) $= 21.48$, $SD = 12.43$; $t(197.01) = 5.95$, $p < .001$, $r = 0.39$ (medium size effect). Closer examination reveals that this performance advantage stems from the fact that students with highest round equal 1 outperform students with highest round equal 1 in the population (M(study) $= 18.7$, $SD = 8.61$; M(population) $= 14.81$, $SD = 10.2$; $t(80.16) = 3.41$, $p < .01$, $r = 0.36$). Students in stage 2 do not differ significantly in competition performance. This supports the conclusions that effects in the general population with regards to cognitive variables might be too conservative in this study. However, all reported effects are interindividual rather than intraindividual, and when disaggregated by different stages (e.g., ordinal regression for stage 1 to 2, and stage 2 to 3) statistical power was not sufficient to detect effected for round 2 to 3 or 4. The reported effects should be taken as interindividual indicators what differentiated students particularly from round 1 and 2. The sample size in the higher rounds is in fact so low that effects cannot separated from noise. Furthermore, general cognitive abilities were assessed through only a subset of items of the full inventory. This has to be factored into conclusions regarding the importance of general cognitive abilities regarding success in the Physics Olympiad.

8. Implications

Implications will be discussed with regards to (1) measuring instruments for evaluating enrichment programs such as the Physics Olympiad, (2) supporting identification and promotion process of students in enrichment programs such as the Physics Olympiad, and (3) empirical status of effectiveness of enrichment programs such as the Physics Olympiad.

Ad (1) Domain-specific skills in evaluation studies for competitions were subsumed under previous competition-related knowledge and conceptualized as prior participation and achievement in the competition. This study factors in an external measure, physics problem solving abilities, such that characteristics of successful participants can be better modelled. Due to the comparatively large proportion of variance that physics problem solving abilities explain in success, it seems advisable to adopt such measures for the other Science Olympiads as well in order to gain a better understanding for the mechanisms that these programs support students. However, physics problem solving abilities is also a fuzzy construct and much finer-grained analyses are necessary in order to outline specific mechanisms that support students thinking skills in physics.

Ad (2) Programs that facilitate engagement for students in these programs are well advised to include elements of problem solving training and meta-cognitive skills on self-knowledge about problem solving. The highest achieving students in the problem solving items presented coherent responses about their approaches to solve these problems. This included conceptual knowledge and also knowledge about applicability of concepts. Thus, these elements should be included in problem solving training – namely outlining an explicit rationale about why something is used in a certain problem context (Fortus, 2009). Furthermore, writing exercises that enable students to reflect their own problem solving were found to be effective. In the literature on meta-cognitive skills and problem solving interventions positive effects for the explicit reflection of strategies were found to be effective (e.g., **???**). Some of the students in this study simply wrote that they have not yet learned the contents that seemed to be required in the problems. Even though this is likely to be true, this response is unfortunate, because modeling physical phenomena and writing about how to tackle the problem is possible without the specific content knowledge to a large extent. This can inform school curricula to adopt more explicit problem solving instruction as a cross-disciplinary skill and anchor this instruction in the respective domain contents (Mayer, 2013; **?**). Furthermore, the positive expectation to be successful in the Physics Olympiad (e.g., expectancy of success) is characteristic of successful students. According to socio-cognitive learning theory feelings of competence and efficacy develop through (amongst others) mastery experiences and positive feedback from meanginful others (**?**). Homework assignments in the competition in the first stage are probably crucial to gain mastery experiences and receive encouraging feedback in order to have a positive expectation towards subsequent stages. Teacher feedback or a welcome letter from the competition could emphasize that the students did a good job and gained mastery experiences that help them solve future problems in physics and thus explaining the value of students' engagement with the problems of the competition for their future lives.

Ad (3) A rather consistent picture evolves considering the effects and effectiveness of Science Olympiads as enrichment programs. Cognitive and affective/motivational variables that are gleaned from the expectancy-value model are viable measures in order to further evaluate these programs. It appears that situation-specific measures (i.e., expectancy of success in Physics Olympiad) are necessary in order to get variance between students. Scales such as self-efficacy in physics or sense of belonging to physics seem too broad (ceiling effects) in order to characterize the participants in these programs (see also: **?**). However, the underlying mechanisms through which these program potentially identify and promote students remain unclear and further longitudinal research is needed to advance this understanding.

Acknowledgement(s)

Disclosure statement

We are not aware of potential conflicts of interest that relate to the results reported in this study.

References

Aljughaiman, A. M., & Ayoub, A. E. A. (2012). The effect of an enrichment program on developing analytical, creative, and practical abilities of elementary gifted students. *Journal of Education for the Gifted*, *35*, 153–174.

Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology*, *30*, 221–256.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W. H. Freeman.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497–529.

Bhaskar, R., & Simon, H. A. (1977). Problem solving in semantically rich domains: An example from engineering thermodynamics. *Cognitive Science*, *1*, 193–215.

Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching*, *43*(5), 485–499.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268.

Docktor, J. L., Dornfeld, J., Frodermann, E., Heller, K., Hsu, L., Jackson, K. A., ... Yang, J. (2016). Assessing student written problem solutions: A problem-solving rubric with application to introductory physics. *Physical Review Physics Education Research*, *12*(1).

Eccles, J. S. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75–138). San Francisco, CA: Freeman and Company.

Feng, A. X., Campbell, J. R., & Verna, M. A. (2001). The talent development of american physics olympians. *Gifted and Talented International*, *16*(2), 108–114.

Fortus, D. (2009). The importance of learning to make assumptions. *Science Education*, *93*(1), 86–108.

Glaser, R. (1992). Expert knowledge and processes of thinking. In D. F. Halpern (Ed.), *Enhancing thinking skills in the sciences and mathematics* (pp. 63–75). Hillsdale, NJ: Erlbaum.

Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, *102*(4), 700–717.

Hammann, M., & Jördens, J. (2014). Offene aufgaben codieren. In D. Krüger, I. Parchmann, & H. Schecker (Eds.), *Methoden in der naturwissenschaftsdidaktischen forschung* (pp. 169–178). Berlin and Heidelberg: Springer Spektrum.

Hazari, Z., Sonnert, G., Sadler, P., & Shanahan, M.-C. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *Journal of Research in Science Teaching*, *47*(8), 978–1003.

Heller, K. A. (Ed.). (2002). *Begabtenförderung im gymnasium: Ergebnisse einer zehnjährigen längsschnittstudie.* Opladen: Leske und Budrich.

Heller, K. A., & Perleth, C. (2007). *Mhbt-s: Münchner hochbegabungstestbatterie für die sekundarstufe.* Göttingen and Bern and Wien: Hogrefe.

Hoffmann, L. (2002). Promoting girls' interest and achievement in physics classes for beginners. *Learning and Instruction*, *12*, 447–465.

KMK. (2009). *Qualitätskriterien für schülerwettbewerbe.*

Köller, O., Daniels, Z., Schnabel, K., & Baumert, J. (2000). Kurswahlen von mädchen

und jungen im fach mathematik: Zur rolle von fachspezifischem selbstkonzept und interesse. *Zeitschrift für Pädagogische Psychologie*, *14*(1), 26–37.
Krapp, A. (2002). Structural and dynamic aspects of interest development: theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, *12*, 383–409.
Larkin, J. H., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*, 1335–1342.
Leonard, W. J., Dufresne, R. J., & Mestre, J. P. (1996). Using qualitative problem–solving strategies to highlight the role of conceptual knowledge in solving problems. *American Journal of Physics*, *64*(12), 1495–1503.
Lubinski, D. (2010). Spatial ability and stem: A sleeping giant for talent identification and development. *Personality and Individual Differences*, *49*(4), 344–351.
Maloney, D. (2011). An overview of physics education research on problem solving: Getting started in physics education research.
Marsh, H. W., Chessor, D., Craven, R., & Roche, L. (1995). The effects of gifted and talented programs on academic self-concept: The big fish strikes again. *American Educational Research Journal*, *32*(2), 285–319.
Mayer, R. (2013). Problem solving. In D. Reisberg (Ed.), *The oxford handbook of cognitive psychology* (pp. 769–778). Oxford and New York: Oxford University Press.
Oswald, F., Hanisch, G., & Hager, G. (2004). *Wettbewerbe und "olympiaden": Impulse zur (selbst)-identifikation von begabungen* (1. Aufl. ed.). Lit.
Petersen, S., & Wulff, P. (2017). The german physics olympiad—identifying and inspiring talents. *European Journal of Physics*, *38*(3), 034005.
Polya, G. (1945). *How to solve it.* New York: Princeton University Press.
Reif, F., & Allen, S. (1992). Cognition for interpreting scientific concepts: A study of acceleration. *Cognition and Instruction*, *9*(1), 1–44.
Reinhold, P., Lind, G., & Friege, G. (1999). Wissenszentriertes problemlösen in physik. *Zeitschrift für Didaktik der Naturwissenschaften*, *5*(1), 41–62.
Reis, S. M., & Renzulli, J. S. (2010). Is there still a need for gifted education? an examination of current research. *Learning and individual differences*, *20*(4), 308–317.
Rheinberg, F., & Vollmeyer, R. (2012). *Motivation.* Stuttgart: Kohlhammer.
Rost, D. H. (2010). *Klare worte zur hochbegabungs-diskussion: Viel seichtes in pädagogik und pädagogischer psychologie.* Retrieved 05.01.2015, from `http://dms-schule.bildung.hessen.de/allgemeines/begabung/Marburger_Hochbegabtenprojekt/HOCHBEGABUNG_F-334R_DIEHL.pdf`
Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology*, *93*, 604–614.
Simon, H. A. (1980). Problem solving and education. In D. T. Tuma & R. Reif (Eds.), *Problem solving and education* (pp. 81–96). Hillsdale, NJ: Erlbaum.
Simon, H. A. (1983). *Reason in human affairs.* Stanford, CA: Stanford University Press.
Snow, R. E., Corno, L., & Jackson, D. N. (1997). Individual differences in affective and conative functions. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 243–310). New York: Simon and Schuster and Macmillan.
Stang, J., Urhahne, D., Nick, S., & Parchmann, I. (2014). Wer kommt weiter? vorhersage der qualifikation zur internationalen biologie- und chemie-olympiade auf

grundlage des leistungsmotivations-modells von eccles. *Zeitschrift für Pädagogische Psychologie*, *28*(3), 105–114.

Subotnik, R. F., Duschl, R., & Selmon, E. H. (1993). Retention and attrition of science talent: A longitudinal study of westinghouse science talent search winners. *International Journal of Science Education*, *15*(1).

Urhahne, D., Ho, L. H., Parchmann, I., & Nick, S. (2012). Attempting to predict success in the qualifying round of the international chemistry olympiad. *High Ability Studies*, *23*(2), 167–182.

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for stem domains: Aligning over fifty years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*, 817–835.

Wai, J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2010). Accomplishment in science, technology, engineering, and mathematics (stem) and its relation to stem educational dose: A 25-year longitudinal study. *Journal of Educational Psychology*, *102*(4), 860–871.

Ziegler, A. (2004). *The actiotope model of giftedness: Ulmer forschungsberichte aus der pädagogischen psychologie* (Nr. 6 ed.).

Ziegler, A., & Stoeger, H. (2009). Identification of underachievement: an empirical study on the agreement among various diagnostic sources. In D. Eyre (Ed.), *Gifted and talented education* (pp. 389–405). London: Routledge.

9. Appendices