Probing the physics problem solving skills of Olympians

P. Wulff[a], S. Petersen[a], K. Neumann[a], and Tim Höffler[a]

[a]Leibniz Institute for Science and Mathematics Education, Olshausenstrasse 62, 24118 Kiel, Germany

ARTICLE HISTORY
Compiled August 5, 2018

ABSTRACT
This template is for authors who are preparing a manuscript for a Taylor & Francis journal using the LaTeX document preparation system and the `interact` class file, which is available via selected journals' home pages on the Taylor & Francis website.

KEYWORDS
Problem solving, Physics competitions

1. Introduction

Modern societies are characterized to be knowledge based and technological (Friedman, 2005). Specialized skills are increasingly important in order to maintain health, wealth, and progress (e.g., Pinker, 2018). In particular, professions in science, technology, engineering, and mathematics (STEM) are relevant in the gamut of required skills in the 21st century. Maintaining health, wealth, and progress is intricately related to understanding STEM skills and training students towards expertise in STEM. However, in times of an aging workforce and a shortage of skilled workers, more talented students need to be recruited for STEM subjects.

As an elaborate means to identify and promote talented students in STEM, enrichment programs such as the science Olympiads were designed (e.g., Petersen & Wulff, 2017). Though research is scarce, some studies suggest that such enrichment programs can be effective in training facets of STEM skills and in developing interest and motivation of students towards STEM (Aljughaiman & Ayoub, A. E. A., 2012; Marsh, Chessor, Craven, & Roche, 1995; Oswald, Hanisch, & Hager, 2004; Wai, Lubinski, Benbow, & Steiger, 2010). Particularly, successful candidates report a positive impact on their future job aspirations in STEM through such programs Feng, Campbell, and Verna (2001); Oswald et al. (2004); Subotnik, Duschl, and Selmon (1993). Other studies report gender differences amongt the particupants in the science Olympiads such as more social support for boys compared to girls (e.g., Lengfelder & Heller, 2002; Urhahne, Ho, Parchmann, & Nick, 2012). Some studies delineate predictors for success in these programs. For example, Urhahne et al. (2012) found for the chemistry Olympiad that previous participation was the best predictor for success in this competition.

However, the empirical status for predicting success in these programs is reminis-

CONTACT P. Wulff. Email: wulff@ipn.uni-kiel.de

cent of Moliere's physician who "explained" opium by its dormitive power (see: Larkin, McDermott, Simon, & Simon, 1980). It just doesn't reach a sufficiently deep level of understanding that would enable researchers to develop targeted interventions that address relevant mechanisms and systematically identify and promote talented students. It would seem necessary to unpack the mechanisms that lead to success further in order to gain knowledge on expertise such that educational institutions can foster the required skills and engage more students with a quality education in STEM. Consequently, this study is designed to move a step further in explaining success and tie success more to relevant cognitive and motivational constructs. As a means to do so, students in the physics Olympiad were tracked throughout their engagement in the physics Olympiad while cognitive and motivational constructs were measured in order to gain a better understanding of factors that predict success in the physics Olympiad.

2. Predictors for success

(Urhahne et al., 2012) identified previous experience, expectancy of success, and general cognitive abilities as predictors for success in the chemistry Olympiad. However, previous experience is a somewhat opaque predictor since a bunch of skills are likely subsumed in this construct which remain unidentified otherwise. Besides motivational constructs such as expectancy of success, we are interested in the cognitive skills that are relevant for success in the physics Olympiad. Overall, an immense amount of experience (years of deliberate practice) in a domain is necessary in order to excel in it (Simon, 1983). We submit that the olympiads tax problem solving as a particularly relevant skill. Problem solving is neither a monolithic skills nor a generic cross-disciplinary ability (e.g., **?**). Problem solving skills have to be assessed within the domain of interest: "It is clear from the problem-solving work done to date, in all domains, that problem-solving expertise is domain-specific" (Maloney, 2011, p. 26)

Expertise in physics problem solving is probably best described with respect to the knowledge that learners bring to problems. Knowledge, relevant for problem solving, can be classified according to declarative and procedural knowledge (Anderson, 1996). Declarative knowledge is organized in chunks, whereas procedural knowledge is organized in production systems that act upon the chunks and can be associated with strategic knowledge. It appears that experts store effective chunks in their memory such that large parts of information in a problem can be processed at once. A chunk is a symbol in long term memory that comprises arbitrary subparts and properties and can be used as a processing unit (Gobet & Simon, 1998). Such chunks can be elicited by production rules (if-then-conditionals) that present learners the conditions for applying the knowledge and present themselves as problem schemata that experts can bring to bear in a certain problem context. Consequently, physics experts are quick to analyze physics problems appropriately. For example, physics experts were found to categorize physics problems based on the relevant principles that are necessary for solving the problems (deep structure), whereas novices categorize problems by surface features such as "is a pulley involved?" (**?**). Deep structure usually is what can be stated verbally in a principle, or in a formula (**?**, p. 70). It is argued that experts in semantically rich domains (e.g., engineering thermodynamics) store a vast amount of domain-specific conceptual knowledge also (Bhaskar & Simon, 1977) that they can match to the correct problems (**?**). Also the knowledge is more hierarchically structured Reinhold, Lind, and Friege (1999). It has been shown that experts in physics problem solving demonstrate more elaborations which are deeper Reinhold et al. (1999). Because experts have a large base of problem schemata (production rules and analogies to solve problems). Based on advantageous knowledge representations, experts solve problems faster and more accurately than novices do (Larkin et al., 1980).

Experts show a careful qualitative problem analysis first, and solve the problem based on the fundamental principles involved (Larkin et al., 1980; **?**). Experts tend to proceed in a logically well-posed sequence of steps of re-translations of the problem, until a certain line of argumentation is reached. Fortus (2009) demonstrates that experts more readily find assumptions for ill-defined problems in order to solve these problems. Novices have problems to recognitize 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.

Besides problem solving abilities in physics, other variables were found to predict success in enrichment programs such as the science Olympiads. The most consistent and important generic personality variable is intelligence. 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). We will focus our discussion on the quantitative and numerical abilities, though we acknowledge that the other dimensions are equally important (Wai, Lubinski, & Benbow, 2009). Even though the correlations of intelligence and complex problem solving seem to depend largely on the utilized constructs (correlations range from .00 to about .80), it seems that particularly well-defined problems (i.e., specific goal-structure) correlate high with facets of intelligence (summarized in: Funke & Frensch, 2007). In general, however, the correlations are rather small (Funke & Frensch, 2007).

For example, Wai et al. (2009) find amongst a sample comprising $400,000$ participants in a 11-year follow-up study 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). However, an interesting phenomenon can be observed towards more expert students. Namely, experts in a STEM field tend to differentiate less by generic abilities (Uttal & Cohen, 2012). This conincides with the general observation that experts rely more on crystallized intelligence (i.e., intelligence that relys on prior knowledge) as compared to fluid intelligence (i.e., intelligence that requires no prior knowledge for problem solving) (Sternberg & Grigorenko, 2003).

Aside from the cognitive factors, also motivational variables were documented to predict success in various aspects in life (**?**). In particular, the expectancy of success in these programs was found to be a relevant factor for predicting success in enrichment programs (Urhahne et al., 2012). In general, the expectation to excel in certain tasks relevant for a domain is an integral part of the achievement-motivation model (**?**). In addition to the expectancy for success the values that are placed upon tasks in the domain are relevant to predict success.

3. Method

In order to identify relevant predictors for success, a host of cognitive and motivational variables were measured that are grounded in the achievement-motivation model (**?**). As motivational constructs we included the expectancy of success in the Olympiad and the values towards the Olympiad as predictors. Another facet was the sense of belongingness to the physics community, which is a relatively novel measure for assessing the motivational affiliation with the competition.

In order to assess quantitative reasoning skills, several subscales of well-established instruments were used K. A. Heller and Perleth (2007). We used three subscales (Q1, Q2, and N1). In particular, N1 involved mental rotation and abstract reasoning, while Q1 and Q2 capitalized on students' understanding of equations and numerical quantities (e.g., what is bigger: $1/2$ or $1/4 + 1/4$). The predictive validity (correlations with school grades one year later) of the N and Q scales are .20, .07, and .45, .35 with science and math, respectively. A subset of items was used such that students from all grades (10,11,12) solved the same items.

Physics problem solving abilities

Of particular importance to us were the physics problem solving abilities of the participants. Maloney (2011) suggests that the nature of the tasks in an investigation (well- versus ill-defined; quantitative versus qualitative), and the definition of what problem solving skills entail should be made explicit. Ad [?]: Four physics problems were presented to the students that are gleaned from two important physics topics: mechanics and E&M. Factual and conceptual knowledge for gravitational and electro-magnetic force, momentum, and energy conservation are necessary. Physics problems that are given to students (e.g., end of chapter-type questions) are typically well-defined problems. Dörner (1976) classified problems according to operations and goal-clarity. Usually, in physics problems, students know the operations that are necessary to solve the problems. The goal-state is also often well-defined and clear.

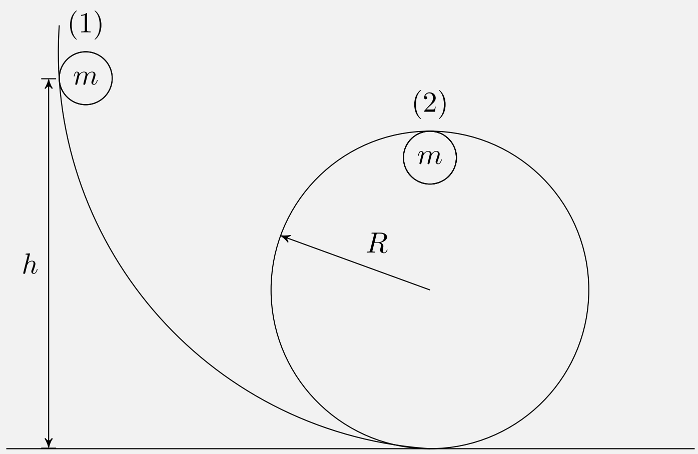
(Leonard, Dufresne, & Mestre, 1996) suggest a measure for physics problem solving that assesses the strategies that students use to solve problems. Strategies, as they use the term, entail the concept to solve the problem, a justification for why the concepts can be applied, and a procedure by which the concept is applied. The authors emphasize the role of conceptual knowledge that is important for physics problem solving, and the aspect that problem solving strategies shall be established before attacking the problem with mathematics. A similar rationale was adopted in the current assessment for problem solving skills. I.e., students were asked to write down their solution strategy for a well-defined physics problem without going through the whole problem solving process (i.e., leaving out problem representation and mathematical routines). The questions were open-ended. In order to make clear to students what was meant with the prompt, an extended example was provided to all students that they read through. Afterwards, four questions followed that each had a unique context (2 mechanics, 2 E&M), but well-defined solutions plans.

The scoring of the open-ended responses utilized measures taken in prior research. Docktor et al. (2016) developed an assessment rubric for scoring students' written solutions for physics problems. This rubric is also useful for the current study, however, not all the five dimensions were meaningful for our measure. Adapting the rubric from the study by Docktor et al. (2016), our scoring rubric comprises the dimensions concept, execution, context, and detail. In the following definitions of the dimensions, we indicate in brackets what the corresponding dimensions from the study by Docktor et al. (2016) are. Concept entails information if students provide all relevant concepts that are necessary for solving the problem ("Physics approach"). Execution entails information about the comprehensiveness of the structure of the solution ("logical progression"). Context provides details about how the concepts are applicable ("specific application of physics"), and detail reflect the assumptions and conditions for applying the concepts ("specific application of physics"). Furthermore, these aspects also relate to the strategy definiton in Leonard et al. (1996). Concept maps to "concept", executions relates to the procedural aspect of strategies, and context and detail relate to the justification aspect of the strategy definition.

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 (e.g., if sentences were logically connected). The answers furthermore differed in the amount of details they provided. Some students interspersed their argumentation with an elaboration of the concepts such as through mathematical formulas. Finally, context accounted for students considerations of boundary conditions in order for them to use the respective concepts. 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).

Example

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.



In addition to the open ended item, the students were also given a test for conceptual knowledge that taxed more the declarative aspect of the knowledge that comes to bearing in solving these problems:

Which equation holds for the mass in point (1) in the picture?

A: $mg = \frac{mv^2}{2R}$
B: $mg + \frac{mv^2}{2R} = 0$
C: $mg = \frac{mv^2}{R}$
D: $mg + \frac{mv^2}{R} = 0$
E: I don't know.

Which equation holds for the mass in point (1) in the picture?

A: $mg2R + mgh = \frac{m}{2}v^2$
B: $mgh = mg2R + \frac{m}{2}v^2$
C: $\frac{m}{2}v^2 = mg2R$
D: $mgh = mg2R$
E: I don't know.

An expert answer for this 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 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) that were acquainted 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 together and adapted when both raters come to an agreement.
Two trained raters scored the students' responses. Agreement was reached in an iterative process where disagreements were discussed and eventually resolved. The final agreement as measured through Cohen's kappa was 0.97.

Instruments

Cognitive measures: Generic physics problem solving is measured with a binary scale (0 incorrect, and 1 correct). The scale contains 15 items (e.g., Welche Tonsaite erzeugt den tiefsten Ton, wenn man sie anzupft?) Content knowledge is measured with a binary scale (0 incorrect, and 1 correct). The scale contains 1 item (e.g., Well-defined physics problem in mechanics and electro-magnetism). Strategy knowledge is measured with an ordinal scale. The scale contains 16 items Cog. abilities is measured with a binary scale (0 incorrect, and 1 correct). The scale contains 1 item. (e.g., What quantity is bigger: $\sqrt{100}$ or 50.)

Motivational measures: Self-efficacy physics is measured with a 4-point Likert scale (not true at all, ..., true). The scale contains 4 items (e.g., Ich bin überzeugt, dass ich alle Fertigkeiten, die zur Lösung von Physikproblemen gebraucht werden, erlernen und beherrschen kann.) Sense of belonging is measured with a 4-point Likert scale (not true at all, ..., true). The scale contains 15 items (e.g., ... vertraue ich darauf, dass die Lehrenden an mich glauben, auch wenn ich etwas schlecht mache.) Expectancy physics competition is measured with a 4-point Likert scale (not true at all, ..., true). The scale contains 5 items (e.g., Ich glaube, dass ich in der PhysikOlympiade erfolgreich werde.)

Success criteria: In order to measure success of the students we either utilized the achievement in round 1 (measured through physics problem for which a maximum of 45 points was awarded). Furthermore, we tracked the highest qualified round (of overall 4) of the students. This gives a proxy for success.

Sample

Table 1. Overview of utilized scales.

Scale	$\bar{N}$	α	$\bar{r}$	Range	M	SD
Generic physics problem solving	119	0.67	0.28	0-1	0.58	0.2
Content knowledge	81.5	0.46	0.21	0-1	0.53	0.29
Strategy knowledge	105	0.87	0.51	-1-2	8.23	6.82
Achievement round 1	124			0-45	28.1	10.11
Highest round	126			1-4	1.71	0.71
Cog. abilities	61	0.86	0.32	0-47	31.1	8.35
Motivational measures:						
Self-efficacy physics	125	0.84	0.68	1-4	3.34	0.55
Sense of belonging	125	0.81	0.42	1-4	3.32	0.34
Expectancy physics competition	125	0.67	0.44	1-4	2.62	0.46
Value physics competition	125.5	0.72	0.52	1-4	2.85	0.62

The sample was drawn from participants in the physics Olympiad. All students that participate in the physics Olympiad were contacted and a fraction of these students took part in the project. Overall the sample considered for this analysis comprised 126 students. These are 45 percent of the whole sample in the project (this smaller number is due to the fact that not all students participated in the second round of the questionnaire where the cognitive measures were assessed). In order to detect possible selection biases, these two samples will be compared with regards to background variables that were administered for both groups. The groups differed with respect to achievement in round 1, M(all) = 25.19, sd = 10.92; M(sub) = 28.1, sd = 10.11; t(255.36) = -2.59, p = .01, r = 0.16. The groups did not differ in terms of highest qualification, M(all) = 1.57, sd = 0.71; M(sub) = 1.71, sd = 0.71; t(238.78) = -1.92, p = .055, r = 0.12, and physics self-efficacy, M(all) = 3.3, sd = 0.39; M(sub) = 3.32, sd = 0.34; t(279.6) = -0.45, p = .65, r = 0.03.

Missing values were not an issue for all scales but cognitive abilities. For cognitive abilities 44 percent are missing. Consequently, we used random regression imputation (Gelman & Hill, 2007) in order to recover the missing values. As calculated via t-test, the imputed scale did not significantly differ from the original scale, M(orig.) = 31.1, sd = 8.35; M(imp.) = 29.53, sd = 8.79; t(151.65) = 1.24, p = .215, r = 0.1. Due to the low internal consistency of the developed content knowledge scale, we omitted this scale from further analyses.

Also, the sample differed from the overall population with regards to achievement, M(sample) = 28.1, sd = 10.11; M(all) = 25.45, sd = 9.05; t(150.46) = 2.77, p = .006, r = 0.22, and with regards to age, M(sample) = 16.29, sd = 1.06; M(all) = 17.1, sd = 1.04; t(159.29) = -8.11, p = 0, r = 0.54.

4. Results

The correlations (Table 2) among the variables indicate that strategy knowledge and generic physics problem solving are significantly related to the highest round of the participants in the competition. However, the motivational scales have no apparent relation to success in the physics Olympiad as measured through achievement in round 1 and highest qualified round. Neither do the motivational measures relate to the other cognitive measures such as content knowledge and strategy knowledge. Moreover, the sense of belonging tends to correlate negatively with the other variables.

The effects of the predictors for success in the physics Olympiad will be analyzed in more detail through a multiple regression. Multiple regression analysis allows to estimate the effects of predictors on a dependent variable while controlling for the other variables in the set. This analysis will give an estimation of relevant predictors for success in the physics Olympiad. We used a stepwise inclusion of predictors, with strategy knowledge entered first in the model because we are interested in the unique contribution of strategy knowledge in particular. Table 4 indicates that strategy knowledge is predictive for success in the physics Olympiad as measured through the highest round achieved and the achievement in round 1. As a second step, we added the covariates into the model. Strategy knowledge remains a significant predictor even after controlling for the covariates such as cognitive abilities. Moreover, expectancy of success in the physics competitions is also a significant predictor for . Furthermore, generic problem solving abilities in physics significantly predict either highest round and achievement in round 1. 21 percent of the variance for highest round, and 14 percent for achievement in round 1 can be explained with the respective model.

Model assumptions and quality: The standardized residuals for the models range from -1.69 to 3.32 for the full model for highest round, and -3.79 to 2 for achievement in round 1. For qualified: No values are of concern related to Cook's distance. No subjects have a concerning leverage. 6 percent of the values have a concerning covariance ratio. For achievement in round 1: No values are of concern related to Cook's distance. No subjects have a concerning leverage. 11 percent of the values have a concerning covariance ratio.

Table 2. Correlations among measured variables.

Measure	1	2	3	4	5	6	7	8	9	10
1 Generic physics problem solving	-	.26	.31*	.33**	.28*	-.02	-.01	.24	-.10	.17
2 Strategy knowledge		-	.33**	.40***	.35**	.09	.000	.18	.19	.30**
3 Achievement round 1			-	.65***	.22	.000	-.08	.25	.09	.05
4 Highest round				-	.26	.11	-.06	.27*	.11	.08
5 Cog. abilities					-	-.03	-.26	.03	.03	.16
6 Self-efficacy physics						-	.38***	.36***	.26	-.17
7 Sense of belonging							-	.33**	.29*	-.01
8 Expectancy physics competition								-	.21	.05
9 Value physics competition									-	-.07
10 Age										-

Table 3. Regression table for simple and full models. Predictor variables are standardized.

	Highest round				Achievement round 1			
	β	$SE(\beta)$	t	p	β	$SE(\beta)$	t	p
Intercept	1.70	0.06	29.12	$< .001$	27.83	0.87	32.13	$< .001$
Prob. sol. strategy	0.18	0.06	2.84	$< .01$	1.78	0.93	1.92	.06
Exp. phy. comp.	0.18	0.06	2.89	$< .001$	2.06	0.92	2.23	$< .05$
Cog. abilities	0.12	0.06	1.93	.06	1.54	0.92	1.67	.10
R^2_{adj}	0.18				0.10			
Intercept	1.70	0.06	29.47	$< .001$	27.86	0.85	32.82	$< .001$
Gen. prob. sol.	0.14	0.07	2.17	$< .05$	2.02	0.96	2.11	$< .05$
Prob. sol. strategy	0.18	0.07	2.75	$< .01$	1.89	0.97	1.94	$< .05$
Cog. abilities	0.07	0.07	1.01	.31	0.70	0.97	0.72	.47
Self eff.	0.02	0.07	0.35	.72	-0.80	1.01	-0.79	.43
Sen. of bel.	-0.10	0.07	-1.42	.16	-1.62	1.00	-1.62	.11
Exp. phy. comp.	0.17	0.07	2.45	$< .05$	2.34	1.02	2.28	$< .05$
Val. phy. comp.	0.01	0.06	0.12	.91	0.46	0.95	0.48	.63
Age	-0.08	0.07	-1.16	.25	-1.46	1.01	-1.45	.15
R^2_{adj}	0.21				0.14			

5. Discussion

In this study we sought to explain success in the physics Olympiad with cognitive and motivational constructs. We were particularly interested in mapping physics problem solving expertise of students to their success in the physics Olympiad in order to gain a better understanding on the functioning of these enrichment programs. In order to assess physics problem solving abilities amongst these high-achieving students, we developed a physics problem solving assessment that would enable us to differentiate students based on their expertise even among this selected sample of participants in the physics Olympiad.

With regards to the physics problem solving scale, strategy knowledge shows satisfying results with regards to internal consistency as measured through Cronbach's α and discriminatory power. Considering content knowledge, these measures are unsatisfatory, especially the internal consistency is low such that this scale was not further included in the analyses. Consequenlty, results with regards to content knowledge should not be interpreted. Strategy knowledge was found to be a relevant variable in order to predict success in the physics Olympiad, as assessed through highest achieved round. However, for achievement in round 1 strategy knowledge only reached marginal significance. Considering achievement in round 1, generic physics problem solving abilities are also a significant predictor. Of the motivational variables, expectancy of success turns out to be a significant predictor for both highest round and achievement in round 1, as was established in prior research. This is well in line with previous findings and the relevant literature (Eccles & Wigfield, 1995; Urhahne et al., 2012).

It should be noted that only a fraction of the students that participated in the physics Olympiad also took part in the questionnaire. The here analyzed sample differed from the overall population with respect to achievement (they perform better), and with regards to age (they are younger). These limitations are important when interpreting the results. However, the importance of expectancy of success can be reinforced in our study. Furthermore, the measure of physics problem solving indicates that the ability to explicate the strategies for the problem solving approach are a facet of what constitutes expertise in the physics Olympiad. Measures that seek to foster physics problem solving expertise among the participants of the physics Olympiad are well advised to reinforce students' qualitative approaches to physics problems.

Acknowledgement(s)

Disclosure statement

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

6. References

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. (1980). *Cognitive psychology and its implications.* New York: Freeman and Company.

Anderson, J. R. (1996). Act: A simple theory of complex cognition. *American Psychologist*, *51*(4), 355–365.

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

de Jong, T., & Ferguson-Hessler, M. G. (1986). Cognitive structures of good and poor novice problem solvers in physics. *Journal of Educational Psychology*, *78*(4), 279–288.

de Jong, T., & Ferguson-Hessler, M. G. (1996). Types and qualities of knowledge. *Educational Psychologist*, *31*(2), 105–113.

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).

Dörner, D. (1976). *Problemlösen als informationsverarbeitung.* Stuttgart: Kohlhammer.

Dreyfus, H. L., & Dreyfus, S. E. (1987). *Künstliche intelligenz.* Reinbek: Rowohlt Taschenbuch Verlag.

Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: the structure of adolescents achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin*, *21*(3), 215–225.

Ericsson, K. A. (2003). The acquisition of expert performance as problem solving: Construction and modification of mediating mechanisms through deliberate practice. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 31–83). Cambridge, UK: Cambridge University Press.

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.

Friedman, T. L. (2005). *The world is flat.* New York: Farrar, Straus, & Giroux.

Funke, J., & Frensch, P. A. (2007). Complex problem solving. In D. H. Jonassen (Ed.), *Learning to solve complex scientific problems* (pp. 25–47). New York: Erlbaum.

Gobet, F., & Simon, H. A. (1998). Expert chess memory: Revisiting the chunking hypothesis. *Memory*, *6*, 225–255.

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

Heller, P., & Hollabaugh, M. (1992). Teaching problem solving through cooperative grouping: Part 2: Designing problems and structuring groups. *American Journal of Physics*, *60*(7), 637–644.

Larkin, J. H., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*, 1335–1342.

Lengfelder, A., & Heller, K. A. (2002). German olympiad studies: Findings from a retrospective evaluation and from in-depth interviews. where have all the gifted females gone? *Journal of Research in Education*, *12*(1), 86–92.

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.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Longman Higher Education.

Ogilvie, C. A. (2009). Changes in students? problem-solving strategies in a course that includes context-rich, multifaceted problems. *Physical Review Special Topics - Physics Education Research*, *5*(2), 139.

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.

Pinker, S. (2018). Counter-enlightenment convictions are 'surprisingly resilient'. *Quillette*.

Pold, J., & Mulvey, P. (2016). *Physics doctorates: Skills used & satisfaction with employment: Data from the degree recipient follow.*

Reif, F. (2008). *Applying cognitive science to education.* Cambridge, MA: MIT 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.

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. (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.

Sternberg, R. J., & Grigorenko, E. L. (Eds.). (2003). *The psychology of abilities, competencies, and expertise.* Cambridge: Cambridge Univ. Press.

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.

Uttal, D. H., & Cohen, C. A. (2012). Spatial abilities and stem education: When, why, and how. *Psychology of Learning and Motivation*, *57*, 147–182.

van Heuvelen, A. (1991). Learning to think like a physicist: A review of research-based instructional strategies. *American Journal of Physics*, *59*, 891–897.

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.

Walsh, L. N., Howard, R. G., & Bowe, B. (2007). Phenomenographic study of students' problem solving approaches in physics. *Physical Review Special Topics - Physics Education Research*, *3*(2).

7. Appendices