Tschisgale *et al.*
Disciplinary and Interdisciplinary Science Education Research (2024) 6:10
https://doi.org/10.1186/s43031-024-00102-y

Disciplinary
and Interdisciplinary Science
Education Research

RESEARCH **Open Access**

Are science competitions meeting their intentions? a case study on affective and cognitive predictors of success in the Physics Olympiad

Paul Leon Tschisgale[1*], Anneke Steegh[1], Stefan Petersen[1], Marcus Kubsch[2], Peter Wulff[3] and Knut Neumann[1]

Abstract
Contemporary science competitions particularly have two intentions: (1) identifying the students demonstrating the highest levels of domain-specific cognitive abilities and (2) recognizing and valuing the efforts of engaged and motivated students, even those without exceptional abilities. This study aimed to examine the relative influence of affective and cognitive variables on predicting success among 136 participants of the first two rounds of the German Physics Olympiad, and based on that, evaluate the extent to which the Physics Olympiad meets the outlined intentions. Our findings indicate that the competition's initial round erects a hurdle for engaged and motivated students who lack sufficient cognitive abilities, which goes against the above mentioned second intention. Conversely, the Physics Olympiad appears to effectively align with its first intention by successfully identifying students with high developed physics-specific abilities. Building on our findings, we discuss ways for better aligning the competition with its intentions, thus contributing to the ongoing further development of science competitions.

Keywor Student competitions, Problem solving, Physics, Expectancy-value model, Logistic regression, Bayesian statistics

Introduction

To enable students to realize their full potential, it is imperative to provide learning opportunities tailored to the students' individual needs (e.g., Smale-Jacobse et al., 2019; U.S. Department of Education, 2013). For students with a strong science affinity, science competitions represent such a learning opportunity. Science competitions challenge students with domain-specific problems to identify the most capable students, while also supporting participating students in further engaging in science and developing science-related abilities (e.g., Abernathy & Vineyard, 2001; Campbell et al., 2000). Quite often, however, science competitions are regarded as tending to perpetuate cognitive elitism, exclusivity, and selectiveness since they are seen to primarily address intellectually gifted students. It remains an objective fact that a crux of science competitions is the identification of students demonstrating the highest levels of domain-specific cognitive abilities. However, another central aim of science competitions is the promotion of all participating students, not just the fraction of exceptionally capable students (e.g., Petersen et al., 2017). Currently,

*Correspondence:
Paul Leon Tschisgale
tschisgale@leibniz-ipn.de
[1]Leibniz Institute for Science and Mathematics Education, Kiel, Germany
[2]Freie Universität Berlin, Berlin, Germany
[3]Heidelberg University of Education, Heidelberg, Germany

science competition managements are actively striving to recalibrate the widespread perception of science competitions as elitist events by actively addressing and supporting a broader student population beyond just the ambit of exceptionally capable students (e.g., Blankenburg et al., 2016; Science Olympiad Inc., 2023). Clearly, this endeavour must go beyond merely increasing the overall number of participating students. In practice, it is of importance that the competition-related efforts of all participating students are recognized and valued (e.g., Avraamidou, 2020). Overall, contemporary science competitions must particularly align two partly contradictory intentions: (1) identifying students demonstrating the highest levels of domain-specific cognitive abilities while also (2) recognizing and valuing the efforts of all participating students.

A large proportion of science competitions consist of multiple rounds, each progressively more challenging, to identify the most capable students. Those participants with the most developed domain-specific cognitive abilities ought to succeed in a specific round and hence advance to the subsequent round. Succeeding at a specific competition round can be regarded as a form of recognition, i.e., successful participants recognize themselves as a competent science person (e.g., Archer et al., 2022) and the competition values their efforts by offering the opportunity to further engage in the competition. While the purpose of the higher competition rounds is in particular to identify the most capable participants, the entry rounds of such multi-round science competitions need closer examination. Typically, such entry rounds aim to encompass a broad range of students. Specifically, there might be engaged and motivated average-ability students who participate and put great effort into the competition. We argue that such students who exhibit beneficial affective attributes (e.g., positive values assigned to the competition, robust self-efficacy beliefs) but lack highly developed cognitive abilities ought to have a reasonable chance of success in the competition's entry round, as a form of recognizing and valuing their efforts. In practice, however, empirical evidence concerning the extent to which science competitions actually (1) successively identify the highest-ability students over the entire course of the competition, while (2) particularly recognizing and valuing the efforts of engaged and motivated average-ability students in the entry round remains somewhat scarce. Given the generally substantial governmental funding of science competitions (e.g., Eremin & Gladilin, 2013; European Commission, 2023), a rigorous evaluation of whether science competitions indeed meet these intentions is essential for the continued development of science competitions as learning opportunities for a broad range of interested students.

An understanding of the extent to which science competitions succeed in both the outlined intentions can be achieved by investigating the relative influence of specific affective and cognitive variables on participants' success (i.e., advancement to the next round) in the entry and subsequent competition rounds. If a science competition inherently succeeds in recognizing and valuing the efforts of engaged and motivated students, affective variables (e.g., values assigned to the competition) ought to have a notable influence on success in a competition's entry round. If a science competition also succeeds in identifying the most capable students, a shift between the first and subsequent competition rounds should be observed in the sense that (domain-specific) cognitive variables become the main driver for success. In short, a better understanding of what contributes to success in the different rounds of science competitions is required. Prior research (e.g., Stang et al., 2014; Urhahne et al., 2012) has started to provide a picture of which variables contribute to success in science competitions. While these studies mainly focused on affective variables, domain-specific cognitive variables can be expected to be particularly predictive of success due to the domain-specific problem solving demands of science competitions. Additionally, most studies generally focused on students' success in single competition rounds only. This way, it remains unclear how the relative role of affective and cognitive variables changes from the entry to subsequent competition rounds, which is– in theory– to be expected if science competitions succeed in their endeavors. Taken together, an in-depth examination of science competitions that uncovers the relative influence of both affective and cognitive variables including domain-specific cognitive abilities on participants' success in the entry and subsequent competition rounds is still pending.

The central aim of the present study was to examine the relative influence of affective and cognitive variables including domain-specific cognitive abilities on success (i.e., advancement) in the first and second round of the German Physics Olympiad– a multi-round science competition for secondary school students (Petersen & Wulff, 2017). By examining what contributes to success in the first and second round of the Physics Olympiad, we can understand to what extent the Physics Olympiad succeeds in (1) identifying the students demonstrating the highest levels of domain-specific cognitive abilities over the first two competition rounds while also (2) recognizing and valuing the efforts of engaged and motivated average-ability students in the entry round.[1] These

[1] The term "success" relates to two different aspects. Given its recurrent use in the manuscript, we want to clarify that we differentiate between (1) success of the Physics Olympiad as an institution in meeting its outlined intentions and (2) success of an individual student in the Physics Olympiad. More precisely, we normatively consider a student successful in a specific round of

findings allow implications on how to improve science competitions and contribute to the continued further development of science competitions.

Theoretical background

The PhysicsOlympiad as a science competition

The national Physics Olympiad in Germany is a science competition for secondary school students that consists of four successive rounds that progressively reduce the number of participants, ultimately revealing the top five students. These top achievers are then invited to represent Germany at the International Physics Olympiad.

In the first round of the German Physics Olympiad, approximately 900 secondary school students voluntarily participate by handing in solutions for the competition tasks. These tasks mainly address standard secondary school physics topics and are solved individually as homework over a period of about five months. Participants succeed in this first competition round and advance to the second round if their scores on the submitted solutions exceed a predefined threshold. On one hand, the competition intends that the most capable students succeed and advance to the second round. On the other hand, it is also intended that the efforts of engaged and motivated average-ability students are recognized and valued (Petersen & Wulff, 2017). The basic structure of the entry round lays the foundation for this intention to be met. More precisely, not only a predefined number of students advances to the next round. This means that engaged and motivated average-ability students have a reasonable chance to be successful in the first round regardless of how many exceptional capable students also advance. The fact that the entry round consists of homework tasks also aligns with the outlined intention as affective student characteristics such as values, self-efficacy beliefs and positive external influences can push a student to engage with the competition tasks over a longer period of time, potentially increasing the student's probability of success. Generally, about 50–70% of the participating students in the first round are then invited to take part in the second round, which consists of tasks that are tackled by the students at home or at school. The tasks of the second and advanced rounds require physics knowledge and abilities exceeding what is typically addressed in regular school curricula. Usually, only half of the qualified students hand in their second round solutions due to the difficulty of the tasks and time constraints in solving them. Of those, approximately the top 50 students are then invited to the third round, in which participants meet each other for the first time in a one-week camp at a research institute. The reason that only a fixed number of students advance is mainly financial (participation is paid in full by the competition, not the students), however, this also aligns with the intention of science competitions to identify the most capable students. Besides theoretical and practical examinations, students are offered opportunities to participate in seminars, excursions, and talks to further develop their motivation and abilities. About 15 of the best students are then invited to the fourth and final competition round whose structure is similar to that of the third round. Finally, the top five students of the fourth round are selected to participate in the International Physics Olympiad. With its multi-round structure and substantial experimental parts in its higher competition rounds, the German Physics Olympiad can be considered largely prototypical among Physics Olympiads worldwide (see Petersen & Wulff, 2017).

the competition if the student advances to the subsequent round, irrespective of this student's individual objectives regarding the competition.

Literature review

Careers of science competition participants

Research on science competitions has extensively investigated how participation in science competitions influences future careers of participating students. Studies by Resch (2013) and Smith et al. (2021) revealed that former science competition participants believed that their participation in the competition had positively influenced their academic and career trajectories. Similarly, Miller et al. (2018) found that, compared to their peers, students who participated in STEM competitions were more likely to engage in a science career, even when controlling for prior STEM interest. Moreover, successful participants in science competitions, i.e., those who performed best at a national level, were found far more likely to perform exceptionally well during their studies and career (e.g., Campbell, 1996; Campbell & Walberg, 2011).

Characteristics of successful participants

The most successful science competition participants, i.e., those students who advanced to the highest rounds in multi-round science competitions, generally engaged in science careers and performed notably above average in their careers. This apparent association appears self-evident taking into account that specific student characteristics can be considered the common cause of both success in the competition as well as subsequent career performance. Therefore, researchers strived to establish an enhanced understanding of successful participants' characteristics, particularly focusing on affective variables. An expectable finding of these studies consisted of successful participants being highly interested in science and in learning about science (Forrester, 2010; Höffler et al., 2019). A retrospective study by Verna and Feng (2002) showed that successful participants generally described themselves as hard-working and being self-disciplined, which these participants considered an important factor of their success. Accordingly, Campbell and Feng (2010) found that less successful participants were often characterized by a lack of motivation. Successful participants were also found to have a high self-concept of ability (Campbell, 1996). This is in accordance with findings of Steegh et al. (2021) who found that the least successful participants in the first round of the German Chemistry Olympiad had the lowest levels of self-efficacy compared to other, more successful participants.

Surveys with former participants revealed that successful participants generally came from families with conducive home atmospheres, e.g., their family members also showed interest in science and supported the students' interests (Campbell & Feng, 2010; Campbell & O'Connor-Petruso, 2008; Verna & Feng, 2002). Steegh et al. (2021) found that the most successful participants in the first round of the German Chemistry Olympiad had experienced the most support from their parents amongst all participants. Moreover, former participants of the German Physics Olympiad attributed a positive influence to their parental home and school (i.e., teacher support, influence of peers; Lind & Friege, 2001). Additionally, successful participants were found to often perceive regular school classes as boring (Verna & Feng, 2002), suggesting that these students were under-challenged by regular schooling.

Next to their effort, successful participants in science competitions generally considered their cognitive abilities as important for their success (Tirri, 2010; Verna & Feng, 2002). Cognitive abilities can be distinguished in general cognitive abilities and domain-specific cognitive abilities. General cognitive abilities refer to more basic abilities (e.g., verbal, quantitative, figural abilities) that are considered largely independent of a domain or subject area (Beauducel & Kersting, 2002). They are generally assessed using measures of intelligence such as IQ tests. One may expect that more successful science competition participants are characterized by higher levels of general cognitive abilities, although research findings are inconsistent in this regard: On one hand, Campbell (1996) found that more successful science competition participants had on average excellent school grades (notably not only in science subjects), which may be indicative of highly developed general cognitive abilities. On the other hand, Lind and Friege (2001) investigated general cognitive abilities of participants in the prefinal and final round of the German Physics Olympiad and found that they were not characterized by particularly high levels of general cognitive abilities, instead they had average abilities. In contrast to general cognitive abilities, domain-specific cognitive abilities are tailored to a specific domain or area, developed through practice and training, leading to an increased performance within that domain or area while potentially having limited applicability beyond it (e.g., Ericsson, 2018). The high domain-specific problem solving demands of science competitions imply that more successful participants ought to be characterized by well-developed domain-specific cognitive abilities which involves profound domain knowledge. Specifically, Campbell and O'Connar-Petruso (2008) reported that successful participants' levels of domain knowledge were far beyond ordinary school knowledge. Moreover, general cognitive abilities were found to predict the initial acquisition of expertise, i.e., acquisition of domain-specific cognitive abilities (Schmidt & Hunter, 2004). Hence, general cognitive abilities matter at the start of expertise development while their role diminishes with increasing expertise and domain-specific cognitive abilities become increasingly important (Ackerman, 1992; Weinert, 2001). In the entry round of science competitions, participants will likely be at different stages in their expertise development. Participants who are at the start of their expertise development may compensate their lack of domain-specific cognitive abilities by well-developed general cognitive abilities. They may therefore still be successful in the first round in which problems typically require a lower level of expertise. In contrast, participants in advanced competition rounds are expected to have developed comparatively high levels of expertise so that their outstanding performance is mostly explained by domain-specific cognitive abilities and only to a much lesser extent by their general cognitive abilities.

Predictors of success

As a result of the plethora of findings on what characterizes successful participants in science competitions, research began to increasingly address the question of what actually determines success in such competitions. Thus, research aimed at determining those characteristics of participants that can be empirically shown to increase the probability of experiencing success in a science competition. These specific characteristics are generally referred to as predictors of success.

At this point, we place a focus on the studies conducted by Urhahne et al. (2012) and Stang et al. (2014) while being aware that further research on predictors of success exists (e.g., Chang & Lin, 2017; Czerniak, 1996; Köhler, 2017). Urhahne et al. (2012) and Stang et al. (2014) adapted the rather broad expectancy-value model of achievement motivation (e.g., Eccles & Wigfield, 2002) to the context of science competitions (see Fig. 1). Urhahne et al. (2012) and Stang et al. (2014) assessed numerous variables from the distinct categories of the adapted expectancy-value model in the German Chemistry and Biology Olympiad. Since they were interested in the total effects of selected predictor variables on performance in the prefinal round in both competitions, they

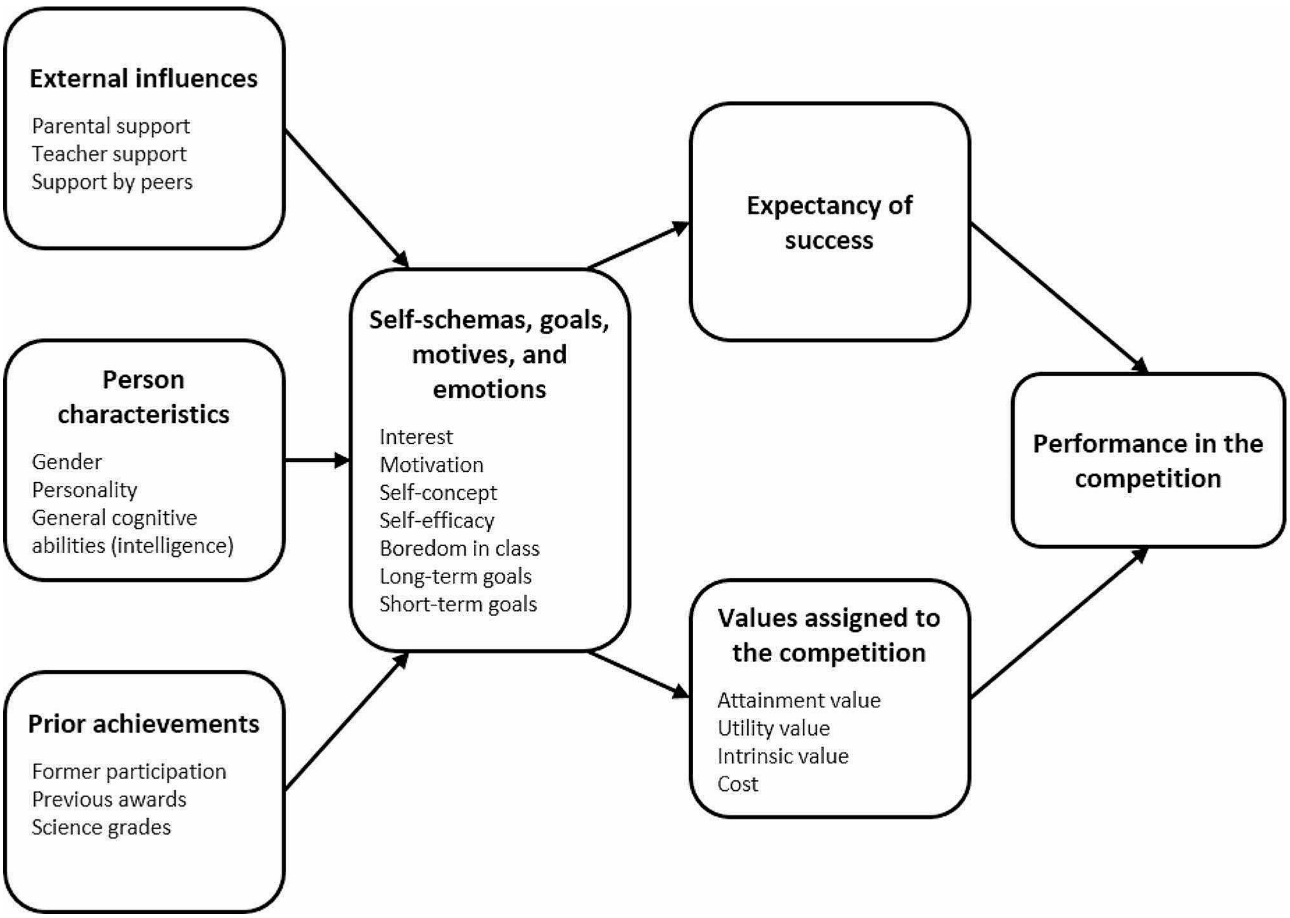

Fig. 1 Expectancy-value model (Eccles et al., 1983) adapted to the context of science competitions

decided to widely ignore the causal structure implied by the expectancy-value model within their analyses.

The underlying idea of their adapted version of the expectancy-value model (see Fig. 1) is that science competition participants' *expectancy of success* and *values assigned to the competition* directly influence their *performance in the competition*. In their model, expectancy of success refers to a student's belief or perception regarding the likelihood of achieving a desired goal which is– in this case– success in the competition. Participants' values assigned to the competition can be expected to influence performance in the competition as value beliefs relate to increased efforts (Guo et al., 2016) generally leading to increased performance. In the model, this value construct is regarded as consisting of four components: Intrinsic value describes to what extent one enjoys participating in the competition combined with one's interest for it (Wang & Degol, 2013). Attainment value reflects the subjective importance of performing well in the competition (Wille et al., 2020). Utility value considers how useful the competition is for the fulfillment of future goals (Shechter et al., 2011). The final component cost subsumes the set of all drawbacks regarding engaging in the competition such as performance anxiety, fear of failure, and expenditure of time (Wigfield & Eccles, 2009).

The two outlined core constructs (i.e., expectancy of success and values assigned to the competition) depend– according to the model– on various other variables that are incorporated within the category *self-schemas, goals, motives, and emotions*. One variable from this category that has been empirically shown to predict science achievement in general (Ferla et al., 2009; Jansen et al., 2015; Parker et al., 2014) and success in science competitions in particular (Steegh et al., 2021) is self-efficacy. Self-efficacy represents an individual's beliefs of being able to successfully perform the necessary actions to reach an anticipated outcome (Bandura, 1977, 1997). While expectancy of success is about the anticipation of positive outcomes, self-efficacy is about one's own beliefs of practically achieving those outcomes, i.e., both constructs differ in their focus.

According to the model, variables in the *self-schemas, goals, motives, and emotions* category are in turn dependent on miscellaneous other variables that are subsumed in the three categories *external influences*, *person characteristics*, and *prior achievements*. Within the category *external influences* students' perceived social support from their parents, teachers, and peers was shown to have a positive effect on achievement outcomes in science education (Cirik, 2015; Ganotice & King, 2014) and also in science competitions (e.g., Campbell & Feng, 2010; Lind & Friege, 2001; Steegh et al., 2021). The category *person characteristics* includes (among others) general cognitive abilities, however, we argue that domain-specific cognitive abilities should also be included in this category due to their outlined importance in science competitions. Specifically, participants in the first round of a science competition will likely be at different stages in their expertise development which is why both general and domain-specific cognitive abilities might be of importance. Problem solving ability represents such a domain-specific cognitive ability.It is regarded as the ability to successfully apply conceptual, conditional, and procedural domain knowledge when dealing with domain-specific problems (Leonard et al., 1996). The category *prior achievements* includes predictors such as participation or achievements in former science competitions, previous awards, and science grades. We argue, however, that using prior achievements to predict future achievements does not actually contribute to a deeper understanding of which student characteristics actually explain success in science competitions because both prior and future achievement can be presumed to have common causes that actually influence success.

Taken together, their version of the expectancy-value model allowed Urhahne et al. (2012) and Stang et al. (2014) to position a wide range of possible predictors of success in science competitions within a single and mature theoretical framework that guided their analyses. In the first step of their analyses, both studies compared successful and unsuccessful participants in the prefinal round (i.e., those who advanced and those who did not advance) based on the pool of assessed variables. As a second step, binary logistic regressions were performed using all significant variables from the first step as predictors for success, i.e., for predicting advancement from the prefinal to the final round in both competitions. Among the wide range of variables under investigation, cognitive variables, however, played a minor role as only nonverbal general cognitive abilities (e.g., visual sequencing and pattern recognition abilities) were considered. Overall, Urhahne et al. (2012) found previous participation in the competition as a significant predictor of success in the Chemistry Olympiad while Stang et al. (2014) found expectancy of success to be a significant predictor of success in the Chemistry Olympiad and perceived boredom in biology classes to be significantly predictive for success in the Biology Olympiad.

The present study

Previous research has started to identify predictors of success in science competitions (Urhahne et al., 2012; Stang et al., 2014). This research, however, has almost exclusively focused on affective variables and general cognitive abilities as predictors, and has typically investigated only a single competition round of multi-round competitions. Specifically, the role of domain-specific cognitive abilities remains underexplored. In consequence, little is known about the relative importance of affective variables, general cognitive abilities, and domain-specific cognitive abilities as predictors of success. Lastly, although the intended focus of science competitions generally changes from entry rounds to advanced rounds from recognizing and valuing the efforts of engaged and motivated students to identifying the students demonstrating the highest levels of domain-specific abilities, research has not addressed to what extent this change of focus corresponds to a change in the relative importance of predictors of success. More precisely, if a science competition meets the intention of recognizing and valuing the efforts of engaged and motivated students, affective variables (in particular values assigned to the competition and self-efficacy) ought to have a notable influence on success in the entry round of the competition. If a science competition also meets the intention of identifying the most capable students, we expect to see an observable shift between the first and subsequent competition rounds in the sense that (domain-specific) cognitive variables become the main driver for success. Thus, in an attempt to better understand what contributes to success in science competitions, we investigated *expectancy of success*, *values assigned to the competition*, *physics self-efficacy*, and *social support* as affective variables and *general cognitive abilities* and *physics problem solving ability* as cognitive variables as predictors of success in the German Physics Olympiad. Specifically, we asked the following research question (see also Fig. 2):

To what extent do both affective and cognitive variables influence the probability of success (i.e., advancement) in the first and second round of the German Physics Olympiad?

Method

Data collection

This study is part of a larger research project (effects of student science competitions, WinnerS) which, among other things, aimed to examine predictors of success and failure in selected science competitions including the German Biology, Chemistry, and Physics Olympiad. This study primarily relies and focusses on data of the Physics Olympiad.

All students who registered for the first round of one of the above competitions including the Physics Olympiad

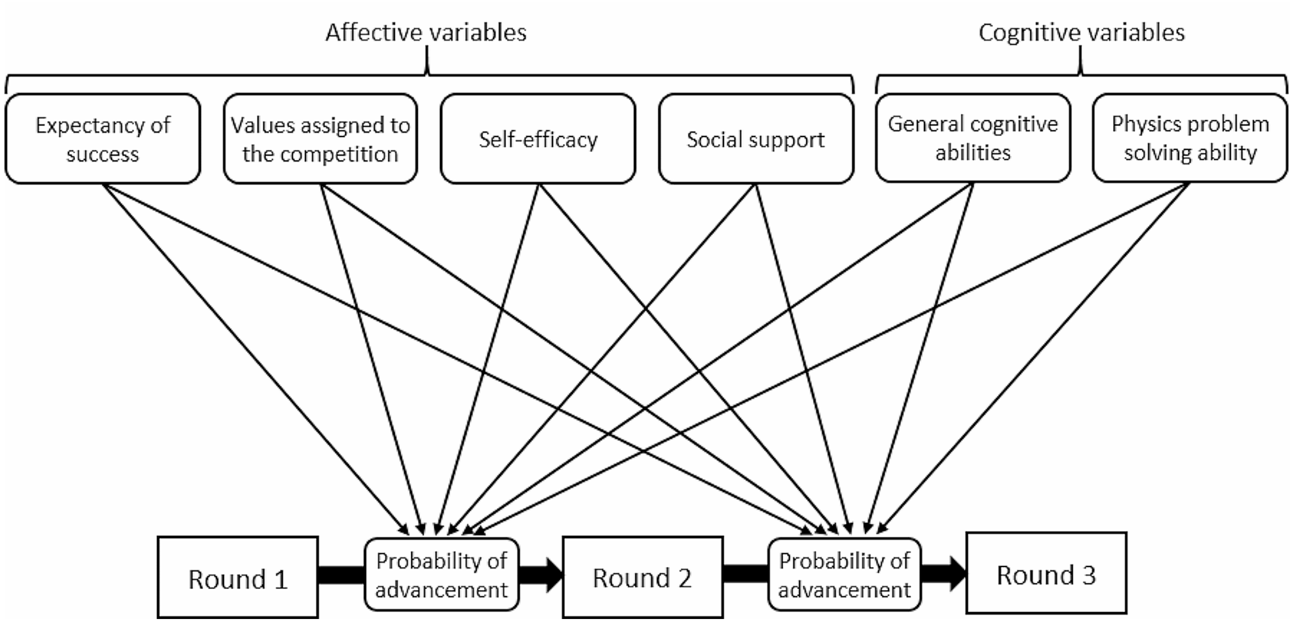

Fig. 2 Theoretical model underlying the research question

received an invitation to voluntarily participate in the study, which consisted of online questionnaires. The study's first questionnaire contained general questions (age, grade, gender, school type) and instruments measuring affective variables. Students had the chance to fill out this first questionnaire until they were informed whether they had advanced to the second competition round. Afterwards, a second questionnaire was unlocked for all students who had participated in the first questionnaire, independent of their success in the first competition round. This second questionnaire focused on measuring cognitive variables. We refrained from assessing both affective and cognitive variables at the same time to prevent participants from terminating the questionnaire before completing it due to massive overload.

Sample

Our study sample consisted of 136 students who participated in the German Physics Olympiad and filled out both the first and the second questionnaire. Of these students, 96.3% attended secondary school academic track (*Gymnasium*). The majority of students in our sample were in 10th (20%), 11th (31%), and 12th grade (43%), while the remaining students were in 8th, 9th or 13th grade (three, two, and three students, resp.).

The selective design of the competition ensured that only a subgroup of participants could advance to the next round. This naturally decreasing number of students reaching the next competition round was also observed in our sample. A comparison between the complete Physics Olympiad population and our study sample with regard to the number of students, their age, and the gender ratio in each of the four competition rounds can be found in Table 1. This comparison indicates that our sample can be considered representative of all Physics Olympiad participants in terms of age and gender ratio up to and including the second competition round.

Table 1 Comparison between the Physics Olympiad population and our study sample for all four competition rounds

	Round 1	Round 2	Round 3	Round 4
Physics Olympiad population				
Number of students	931	406	53	15
Mean age (*SD*)	16.3 (1.1)	16.3 (1.0)	16.7 (0.8)	16.5 (0.8)
Gender ratio (male/female)	0.72/0.28	0.73/0.27	0.89/0.11	0.93/0.07
Study sample				
Number of students	136	77	14	3
Mean age (*SD*)	16.2 (1.1)	16.3 (1.0)	16.5 (0.5)	16.3 (0.6)
Gender ratio (male/female)	0.70/0.30	0.74/0.26	0.79/0.21	0.67/0.33

Note All participants in the competition and therefore in our sample identified as either male or female

Instruments

Affective predictor variables

The following instruments measured affective variables and participants were asked to specify their agreement to given statements on 4-point Likert scales ranging from "I completely disagree" (1) to "I completely agree" (4). We decided for an even-numbered Likert scale to prevent possible mid-point bias (Garland, 1991).

Expectancy of success. This construct was measured with the four items "I believe that I will be successful in the Physics Olympiad", "I imagine that I will have problems learning what I have to in the Physics Olympiad", "I expect to do better than many other Physics Olympiad participants", and "I think I can acquire the knowledge I need for the Physics Olympiad". These items were selected from scales by both Eccles and Wigfield (1995) and Lykkegaard and Ulriksen (2016) and adapted in order to relate to the Physics Olympiad. As the original scales relate to mathematics and STEM study programs, we do not expect the adaptions to have any influence on individual item validity. Moreover, internal consistency of the scale in terms of Cronbach's alpha as an estimate of reliability proved acceptable ($\alpha = 0.71$).

Values assigned to the competition. To measure this construct, we used the scale developed by Lykkegaard and Ulriksen (2016). Specifically, we decided to use a single item to measure each of the four existing value components for test-economic reasons. The wording of the four items was adapted to conform to the Physics Olympiad. The used items read: "I get involved in the Physics Olympiad because I find it very interesting" (intrinsic value), "It means a lot to me to be good in the Physics Olympiad" (attainment value), "I expect that what I will learn in the Physics Olympiad will also be beneficial in my everyday life" (utility value), and "It is important for me to get involved in the Physics Olympiad even if I will have less time for family, friends and leisure activities" (cost). Again, as the original scale relates to STEM study programs in general, we do not expect our adaptions to impact individual item validity. Moreover, internal consistency of the scale proved acceptable ($\alpha = 0.75$).

Self-efficacy. Physics self-efficacy was measured with an adapted version of the complete mathematics self-efficacy scale from the German national questionnaire of the PISA studies (PISA-Konsortium Deutschland, 2006). The original scale consists of four items that were all adapted by replacing the word "mathematics" with "physics" which we do not expect to have any impact on construct validity. The used items read: "I am confident to understand even the most difficult material in physics", "I am convinced that I can solve even the most complicated physics tasks", "I am convinced that I can always achieve very good results in physics", and "I am convinced that I can learn and master all abilities needed to solve physics problems." The internal consistency of the scale proved good ($\alpha = 0.85$).

Social support. To measure participants' perceived social support with regard to physics and the Physics Olympiad, we combined the 'support by parents', 'support by teachers', and 'support by peers' scales developed by Wulff et al. (2018). Three items from each scale were used whereby items have a similar structure across the original scales: "My parents/teacher/friends supported me very much regarding the Physics Olympiad", "I can turn to my parents/teacher/friends if I have problems in or questions about physics", and "My parents/teacher/friends actively support(s) me in my physics engagement." The internal consistency of the combined scale was acceptable ($\alpha = .72$).

Cognitive predictor variables

General cognitive abilities. In order to assess general cognitive abilities, we used a subscale of a cognitive abilities test developed by Heller and Perleth (2007) in which students receive different items according to their grade level. Specifically, we chose the subscale for quantitative cognitive abilities as quantitative abilities are of central importance in science (e.g., Wai et al., 2009) and in particular in the Physics Olympiad (Treiber et al., 2023). A sample item from this subscale reads "Which quantity is bigger: q^3 or q^4 if q is real and a positive proper fraction?"

Physics problem solving ability as a domain-specific ability. Existing instruments for assessing this ability (e.g., Brandenburger, 2016; Coleman & Shore, 1991) were not designed for particularly capable students as those that can be found in the Physics Olympiad, and are therefore at risk of exhibiting ceiling effects. Hence, we designed a new instrument to measure physics problem solving ability focussing on students' strategies for solving a given problem (to find the complete instrument, see Wulff et al., 2023). Such strategies entail the concepts to solve a given problem (conceptual knowledge), a justification for why these concepts can be applied (conditional knowledge), and procedures by which these concepts are applied (procedural knowledge). The designed instrument requires students to describe in written form and in full sentences how they would solve four well-defined physics problems without explicitly solving them. A theory-based coding rubric distinguishing the four categories *concept*, *context*, *execution*, and *detail* was used for scoring students' responses to each problem. All responses were completely double coded by two raters. Initial agreements in the four categories measured through Cohen's linearly weighted kappa (Warrens, 2012) were substantial to almost perfect ($\kappa_{\text{concept}} = 0.81$, $\kappa_{\text{context}} = 0.85$, $\kappa_{\text{execution}} = 0.77$, $\kappa_{\text{detail}} = 0.79$; Landis & Koch, 1977). In order to further increase the quality of the ratings, disagreements between raters were discussed until a consensus was reached.

Success in the competition

Success in a specific competition round was considered a dichotomous variable, i.e., a student was successful in a specific round if this student advanced to the subsequent round and vice versa. This decision on advancement, in turn, was based on the scores on participants' submitted solutions in the corresponding round. More specifically, participants needed at least 30 of 40 points in the first round in order to advance to the second round, while roughly 50 students with the highest scores among all second-round participants advanced to the third round.

Analyses

The central aim of this study was to examine the influence of *expectancy of success, values assigned to the competition, self-efficacy, social support, general cognitive abilities* and *physics problem solving ability* as affective and cognitive predictor variables on advancement in the first and second round of the German Physics Olympiad (see Fig. 2). More specifically, advancement in a specific competition round was regarded as a dichotomous outcome as it was operationalized by whether a participant advanced to the next competition round or not. Taking these aspects into account, we applied logistic regression analyses as they allow the analysis of the effect of multiple independent variables (the aforementioned affective and cognitive predictor variables) on a dichotomous outcome variable (success in a specific round) by quantifying each independent variable's unique contribution (Stoltzfus, 2011). To prepare data for these logistic regression analyses, we performed two preliminary analyses: a Rasch analysis of the general cognitive abilities data and a multiple imputation procedure to handle missing data. All statistical analyses in this study were conducted using R (Version 1.4; R Core Team, 2021).

Rasch analysis of general cognitive abilities data

In contrast to the other instruments used in this study, participants did not receive the exact same items in the test for general cognitive abilities. Rather, participants received a grade level-specific subset of items. Since there existed common items between subsets, Rasch modelling would allow expressing each participant's general cognitive ability as a score on the same scale irrespective of which grade level-specific subset of items was answered by a participant. However, this only holds if the relevant construct is unidimensional (Boone & Noltemeyer, 2017) which should be the case since we have only used the subscale for quantitative abilities of the entire instrument measuring general cognitive abilities. Thus, in order to obtain comparable values representing general cognitive abilities across all participants, we performed a Rasch analysis using the R package *TAM* (Robitzsch et al., 2021). The basic idea of a Rasch model and analysis is that the probability of a person correctly solving a specific item only depends on the difference between the person's ability score and the specific item's difficulty score. This way, a Rasch analysis constructs ability scores which are on the same scale irrespective of the specific items answered by the participants (Boone & Noltemeyer, 2017; DeMars, 2010).

Before estimating the Rasch model, items that all participants answered correctly or incorrectly must be removed. Then the model is estimated and the model results are inspected in an iterative process, i.e., the inspection of model results may lead to the exclusion of specific items based on criteria and to a re-estimation of the model. Specifically, we focused on three criteria. First, each item's infit value which represents a measure of fit between the item and the Rasch model should be located between 0.8 and 1.2 (Bond & Fox, 2007). Second, we inspected Wright Maps that contrast estimated person ability scores and item difficulty scores in order to illustrate whether the set of items covers the whole set of abilities of participants (Bond & Fox, 2007). Third, we computed the weighted likelihood estimate (WLE) reliability which represents an overall measure of fit between the data and the Rasch model and which can be interpreted similarly to Cronbach's alpha (Adams, 2005).

Since only 75 students from our overall sample ($N=136$) completed the cognitive abilities test, we may have obtained unreliable estimators for students' cognitive ability scores due to a too small sample size (Neumann, 2014). Hence, to increase sample size for estimation, we included Biology and Chemistry Olympiad participants' test results, which were assessed in the same project. This provided us with a total sample of 495 students for estimating the Rasch model.

Multiple imputation for missing data

As in other survey-based empirical research, we faced the issue of missing data. Specifically, only 75 students from our overall sample ($N=136$) participated in the test for general cognitive abilities. This was likely caused by the length of the overall test procedure as the test for general cognitive abilities was to be immediately processed after the assessment of physics problem solving abilities. In order to address this issue, we used multiple imputation as a recommended method (Enders, 2010). A common criticism is that multiple imputation cannot handle large amounts of missingness. However, traditional methods such as listwise deletion would generally be inappropriate in such situations as they introduce bias and largely reduce statistical power (van Ginkel et al., 2020). Moreover, simulation studies showed that multiple imputation can handle even large missing rates (e.g., Grund et al., 2016; Madley-Dowd et al., 2019).

Multiple imputation is a regression-based procedure which consists of three steps (van Ginkel et al., 2020): First, multiple complete copies of the incomplete data are generated by replacing the missing values with different plausible estimates. Second, all of these complete versions are analysed separately by the intended statistical procedure, which will result in slightly varying outcomes of the analysis. Third and finally, these slightly varying outcomes are combined into a final result by an appropriate statistical procedure which takes into account the uncertainty induced by the missing data.

In our study, we used the R package *mice* (van Buuren & Groothuis-Oudshoorn, 2011) to perform this multiple imputation procedure while also following recommendations by Zhou and Reiter (2010). That is, we created $m=100$ complete versions of the incomplete dataset to ensure reliable inferences (first step). Technical details regarding this step can be found in the Supplementary Material (Part B). Moreover, we scaled all predictor variables in each complete dataset ($M=0$, $SD=0.5$) to ease future interpretation and comparison of results. Logistic regressions were then performed on all $m=100$ complete dataset (second step) and finally all analysis outcomes were combined into a final result (third step).

Bayesian logistic regression to answer the research question

In order to investigate the effects of the introduced affective (expectancy of success, values assigned to the competition, self-efficacy, and social support) and cognitive predictor variables (general cognitive abilities and physics problem solving ability) on the probability of advancement in the first and second round of the German Physics Olympiad, we performed two logistic regressions– one for each transition between competition rounds. A logistic regression model provides us with a regression parameter for each predictor variable that describes the strength and direction of the influence of that predictor variable on an outcome. Hence, in our context, we will obtain estimates (in the form of regression parameters) that describe the individual influence of the six introduced predictor variables on the probability of advancement in either the first or second round of the Physics Olympiad. Such logistic regression models were previously successfully applied in similar studies (Urhahne et al., 2012; Stang et al., 2014). To estimate the regression models, we chose the R package *brms* (Bürkner, 2017) which incorporates an efficient way to handle multiple imputed datasets as it uses a Bayesian approach for model fitting. Further information regarding the specification of our logistic regression models and their estimation can be found in the Supplementary Material (Part C).

This Bayesian approach for model fitting does not provide us a single estimate for a desired regression parameter as in the frequentist approach, but rather a probability distribution over all possible values of that regression parameter. This distribution is referred to as posterior distribution and does not only contain information on the most probable value of a regression parameter but also its level of uncertainty. More precisely, this most probable value of a regression parameter is given by the maximum a posteriori (*MAP*) estimate which corresponds to the mode of this parameter's posterior distribution. The *MAP* estimate can be considered the Bayesian counterpart of the traditional point estimate in the frequentist approach. The uncertainty of this *MAP* estimate is quantified by the 95% highest posterior density interval (*HPDI*) which is the narrowest interval of a posterior distribution containing the specified probability mass (here 95%). The *HPDI* can be considered the Bayesian counterpart of the traditional confidence interval in the frequentist approach. In summary, we will provide both *MAP* estimates and their corresponding 95% *HPDIs* to describe regression parameters when presenting our results. A reader more interested in Bayesian methods may see Kubsch et al. (2021) for an introduction in the context of science education research or McElreath (2020) for an in-detail treatment of the subject.

Moreover, the posterior distributions of regression parameters can be used to compute probabilities of advancement to the next competition round based on the effects of single predictor variables. This makes it possible to make exact statements about the extent to which a change in a predictor variable influences the probability of advancement to the next round. In particular, these changes in probability are much easier to interpret than concrete values of regression parameters.

Results

Descriptive statistics

Means, standard deviations, and correlations of unscaled predictor variables are presented in Table 2. The statistically significant correlations are all positive, have small to moderate magnitudes, and occur either between the affective variables or between the two cognitive variables.

Preliminary analyses

In order to prepare the general cognitive abilities data for further analyses, we performed a Rasch analysis. The final Rasch model showed a satisfying fit with acceptable infit values between 0.8 and 1.2 (Bond & Fox, 2007) and a WLE reliability of 0.77. Moreover, the Wright Map revealed that the items of the test for general cognitive abilities cover the whole range of ability levels of participants in an acceptable way (see Supplementary Material, Part A). Thus, we obtained reliable estimates of participants' general cognitive abilities that were used for further analyses.

Moreover, we handled missing data by performing multiple imputation as a recommended method (Enders, 2010). Graphical diagnostics using time-series plots indicated that the imputation method produced reliable estimates to replace missing values (see Supplementary Material, Part B).

Table 2 Means, standard deviations, and correlations of unscaled predictor variables

Variables	M	SD	EXSU	VACO	SEEF	SOSU	GCAB
Affective variables							
Expectancy of success (EXSU)	2.56	0.48					
Values assigned to the competition (VACO)	2.86	0.63	0.19*				
Self-efficacy (SEEF)	3.32	0.57	0.32***	0.25**			
Social support (SOSU)	2.52	0.52	0.17	0.04	0.01		
Cognitive variables							
General cognitive abilities (GCAB)	0.65	0.88	0.20	0.01	0.16	−0.15	
Physics problem solving ability (PPSA)	7.57	6.74	0.12	0.15	0.09	0.03	0.38***

Note M=mean; *SD*=standard deviation

*$p \leq .05$

**$p \leq .01$

***$p \leq .001$. The statistics of the *general cognitive abilities* data correspond to the Rasch-modelled person abilities

Table 3 Effects of predictor variables on advancement in the first and second competition round

	Advancement from R1 to R2		**Advancement from R2 to R3**	
Variables	*MAP*	95% *HPDI*	*MAP*	95% *HPDI*
Affective variables				
Expectancy of success (EXSU)	0.94	[0.08, 1.83]	-0.37	[-1.85, 1.03]
Values assigned to the competition (VACO)	-0.11	[-0.91, 0.70]	0.52	[-0.83, 1.95]
Self-efficacy (SEEF)	0.02	[-0.82, 0.83]	0.97	[-0.64, 2.70]
Social support (SOSU)	0.11	[-0.67, 0.96]	0.16	[-1.18, 1.46]
Cognitive variables				
General cognitive abilities (GCAB)	0.91	[-0.24, 2.15]	0.00	[-1.92, 2.17]
Physics problem solving ability (PPSA)	1.59	[0.47, 2.54]	1.01	[-0.18, 2.41]

Note R1=first competition round; R2=second competition round; R3=third competition round; *MAP*=maximum a posteriori estimate; 95% *HPDI*=95% highest posterior density interval

Logistic regression analyses to answer the research question

The results of the logistic regression analyses for the effects of predictor variables on the probability of advancement in the first and second competition round in the form of *MAP* estimates and corresponding 95% *HPDI* are shown in Table 3. Complete posterior distributions of estimated regression parameters can be found in the Supplementary Material (Part D).

Moreover, we determined probabilities of advancement to the next competition round based on effects of single predictor variables (see Fig. 3). The less the depicted curves overlap for a particular predictor variable, the greater the influence of that variable on success. If, on the other hand, the curves overlap almost completely, the corresponding predictor variable has almost no influence on success.

We found expectancy of success and both cognitive variables to have a notable influence on the probability of advancement from the first to the second competition round. Physics problem solving ability was the strongest predictor of success (MAP_{PPSA} = 1.59), followed by expectancy of success and general cognitive abilities with a comparable influence on success (MAP_{EXSU} = 0.94, MAP_{GCAB} = 0.91). By examining the left side of Fig. 3, we can clearly recognise these strong influences on the probability of advancement to the second round. A participant with an average value on each predictor variable has an average probability of advancement of approximately 60%. If, on one hand, an average participant's physics problem solving ability increases by one standard deviation, this participant's average probability of advancement will increase by about 17%. One the other hand, a decrease of one standard deviation would result in a decrease of the average probability of advancement by 20%. A similar change of an average participant's expectancy of success or general cognitive abilities would change the average probability of advancement by 12% in the appropriate directions. Compared to these three variables, the remaining predictor variables (values assigned to the competition, self-efficacy, and social support) had a negligible influence on advancement (MAP_{VACO} = -0.11, MAP_{SEEF} = 0.02, MAP_{SOSU} = 0.11). This can also be observed in Fig. 3, as there is nearly no change in the probability of advancement to the second round when considering different values of the corresponding predictor variables.

A different picture emerges for advancement from the second to third competition round. One notable observation is that the range of the *HPDI* has increased compared to those of the first round, which indicates a higher uncertainty about regression parameter values. This is mainly attributed to the smaller sample size for estimation as there are naturally fewer participants in the second round compared to the first round (see Table 1). Comparing all *MAP* estimates, we found self-efficacy

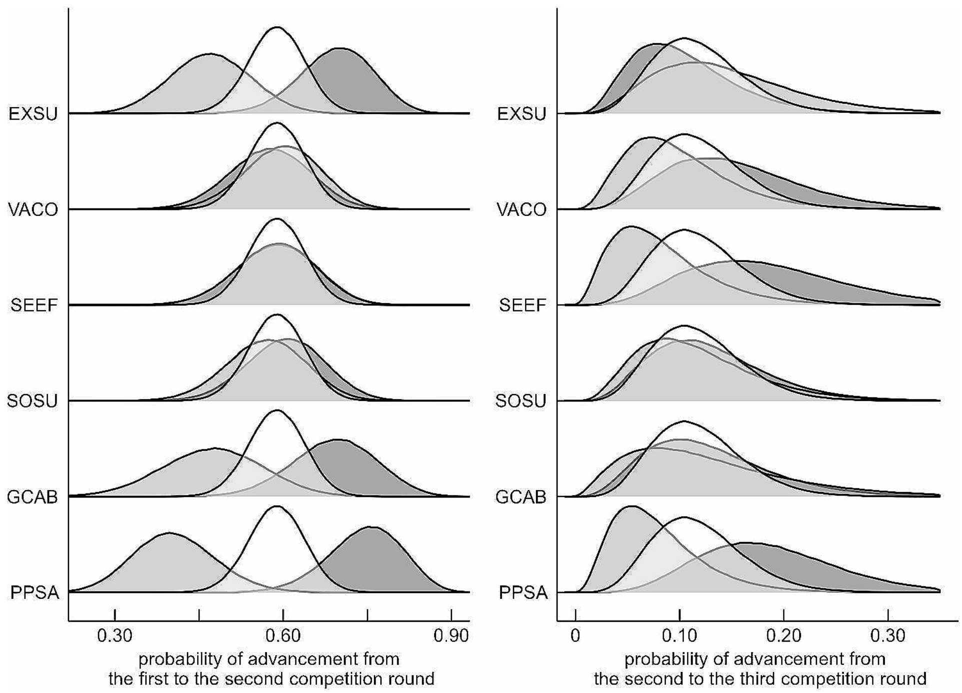

Fig. 3 Probabilities of advancement to the next competition round for different values of predictor variables *Note.* EXSU = expectancy of success; VACO = values assigned to the competition; SEEF = self-efficacy; SOSU = social support; GCAB = general cognitive abilities; PPSA = physics problem solving ability. Regarding one transition of interest, three differently shaded curves are shown for each predictor variable. The middle curve (transparent) represents the probability of advancement to the next round of participants who have an average value on each predictor variable (here *average* refers to the participants of our study). The light grey and dark grey shaded curves correspond to participants who have a low (one standard deviation below the average) and high value (one standard deviation above the average) on the corresponding predictor variables, respectively, while all other variables are kept at average level

and values assigned to the competition amongst the affective variables and physics problem solving ability amongst the cognitive variables to have a notable influence on the probability of advancement from the second to third competition round. In particular, physics problem solving ability can once more be considered the best predictor of success, this time in the second competition round (MAP_{PPSA} = 1.01). However, self-efficacy is nearly as strong a predictor (MAP_{SEEF} = 0.97), while values assigned to the competition seem to have a moderate influence on the probability of success as well (MAP_{VACO} = 0.52). Again, by examining the right side of Fig. 3, we can observe the effects of these predictors on the probability of advancement to the next round. First, an average participant has an average probability of advancement of only 11%, which goes hand in hand with the stronger selective character of the second competition round. Increasing or decreasing an average participant's problem solving ability by one standard deviation would change this participant's average probability of advancement by 6% in the appropriate directions. Similar considerations for self-efficacy and values assigned to the competition result in a change in the probability of advancement by 5% and 3%, respectively, in the appropriate directions. The remaining predictors (expectancy of success, social support, and general cognitive abilities) seem to have no noticeable influence on the probability of advancement in the second competition round (MAP_{EXSU} = -0.37, MAP_{SOSU} = 0.16, MAP_{GCAB} = 0.00) which can also be concluded by examining Fig. 3.

Discussion

This study aimed to understand the extent to which the Physics Olympiad succeeds in reconciling its intentions of (1) identifying the most capable students and (2) recognizing and valuing the efforts of engaged and motivated average-ability students. For this purpose, the present study examined the relative influence of affective and cognitive variables including domain-specific cognitive abilities on success in the first and second round of the German Physics Olympiad. If the Physics Olympiad meets the intention of recognizing and valuing the efforts of engaged and motivated students, then affective variables ought to have a notable influence on success in the entry round of the competition. If the Physics Olympiad also meets the intention of identifying the most capable students, then there should be an observable shift between the first and subsequent competition rounds in the sense that (domain-specific) cognitive variables become the main driver for success. More specifically, this study performed logistic regression analyses to quantify the relative effects of the predictor variables *expectancy of success, values assigned to the competition, self-efficacy, social support, general cognitive abilities* and *physics problem solving ability* on the probability of success in the first and second competition round.

Predictors of success

Advancement from the first to second round

We found that certain variables are notably related to an increased probability of success in the first competition round, i.e., advancing to the second competition round. Physics problem solving ability had the most notable effect on success in the first competition round. This observation aligns with insights from expertise research which highlights that domain-specific cognitive abilities acquired through deliberate practice and experience substantially contribute to outstanding performance within a given domain (Ericsson, 2018). However, we also found general cognitive abilities to play a notable role in the first round of the Physics Olympiad. This finding suggests that even participants who were in the early stages of their expertise development and therefore primarily relied on their general cognitive abilities (Lind & Friege, 2001; Weinert, 2001) had a reasonable chance of succeeding in the first competition round. More precisely, quantitative abilities as a specific facet of general cognitive abilities were assessed, which is why well-developed quantitative abilities seem to be important for success in the entry round of the Physics Olympiad, a finding that aligns with the conclusions of Treiber et al. (2023).

Participants' expectancies of success were also found to have a notable influence on success. This finding aligns with previous research which demonstrated a positive correlation between expectancies of success and students' achievements (Guo et al., 2016; Trautwein et al., 2012), particular in the context of science competitions (Stang et al., 2014). However, this finding was somewhat predictable considering that participants rated their expectancy of success during or after engaging with the first round's tasks. Consequently, these personal ratings were specifically linked to these tasks. Hence, this finding may be seen as an indicator that participants were particularly good at predicting their own performance on the first round's tasks.

Among the other affective variables examined (i.e., values assigned to the competition, self-efficacy, and social support), none exhibited a notable influence on success. The finding that social support did not have a notable effect on success contradicted our initial anticipation. We hypothesized results similar to those of Steegh et al. (2021) who identified participant profiles in the Chemistry Olympiad and found that students in the most successful profile received the most parental support. Simpkins et al. (2015) found that parental support predicted adolescents' science-related self-efficacy and values which in turn influenced academic success. However, our data suggests neither a direct nor an indirect effect (mediated through self-efficacy or values assigned to the competition) of social support on success since we did not find any significant correlation between social support and the two possible mediators (see Table 2). Although social support seems to play no role in explaining success in the first round of the Physics Olympiad, it could still potentially explain the decision to participate in the competition in the first place (Czerniak, 1996; Verna & Feng, 2002).

Based on our findings, we concluded that the Physics Olympiad does not succeed in meeting the intention of inherently recognizing and valuing the efforts of engaged and motivated average-ability students since neither values assigned to the competition nor self-efficacy nor social support were found to have a notable influence on success in the entry round of the competition. Specifically, general cognitive abilities and domain-specific cognitive abilities in the form of physics problem solving abilities were found to increase the probability of success in the first competition round the most. Overall, our findings indicate that a lack of cognitive abilities cannot be compensated by highly developed affective variables. Nonetheless, our findings suggest that less developed physics problem solving abilities may be compensated by well-developed general cognitive abilities– and vice versa– as both were found to notably increase the probability of success in the first round. Taken together, success in the first round of the German Physics Olympiad requires more than engagement and motivation. It seems that already in the first round successful students possess highly developed physics problem solving abilities or are able to compensate a lack of those domain-specific cognitive abilities by well-developed general cognitive abilities.

Advancement from the second to third round

We also examined the relative influence of affective and cognitive variables on success in the second round of the Physics Olympiad. This allowed for contrasting the relative contribution of predictor variables on success and compare it to the intended shift of focus from recognizing and valuing the efforts of engaged and motivated students to identifying the most capable students.

In contrast to the findings of the first round, participants' expectancy of success had no notable effect on success in the second competition round. This finding, however, may be a consequence of a methodological issue. We assessed participants' expectancies of success during the first competition round using items which explicitly addressed a general expectancy of success concerning the competition as a whole, rather than focussing on a specific round. Yet, we suspect that a majority of participants based their expectancy beliefs on their experiences in the tasks of the first round. Moreover, tasks of the second round differ from those of the first round as they are more challenging (Petersen & Wulff, 2017). Thus, the estimated effect of participants' expectancies of success on advancement in the second competition round must be interpreted with caution.

Participants' values assigned to the competition were found to have no influence on success in the first round, but had a notable influence in the second round. As the tasks of the second round are more difficult and particularly more time-consuming than tasks of the first round (Petersen & Wulff, 2017), participants generally have to show more commitment and effort to solve the tasks within the given timeframe. Given that value beliefs have been established as being related to students' efforts (Guo et al., 2016), this might explain why students' values contribute notably to success in the more demanding second round of the competition.

Self-efficacy, which exhibited no notable influence on success in the first round, emerged as the second-strongest predictor of success in the second round. This shift could be attributed, once more, to the increased difficulty of competition tasks in the second round. Considering that physics self-efficacy represents participants' beliefs of being able to understand even the most difficult physics material and successfully tackle the most challenging physics problems (Bandura, 1977, 1997), participants possessing high self-efficacy would be more successful in the second round when compared to their counterparts with lower self-efficacy.

General cognitive abilities were found to have no influence on success in the second round even though they had a notable influence in the first competition round. In contrast, physics problem solving ability remained the best predictor of success even in the second round. The importance of this finding becomes all the more clear from a statistical point of view. Those participants who advanced to the second round generally had better developed problem solving abilities than average participants of the first round since problem solving abilities were found to be most predictive of success in the first round. Therefore, a reduced variance of participants' physics problem solving ability could be expected in the second round as participants with less developed physics problem solving abilities were less likely to advance to the second round. Despite this reduced variance, physics problem solving ability remained the best predictor of success in the second round. This highlights that the Physics Olympiad seems to identify students with well-developed general cognitive abilities or problem solving abilities in the first round, and students with in particular even better developed physics problem solving abilities in the second round.

These findings regarding both general cognitive abilities and physics problem solving ability as a physics-specific ability can be linked to expertise research. The notable role of general cognitive abilities in the first round may indicate that even participants, who were still at an early stage of their expertise development in physics (Ericsson, 2018), had a reasonable chance to succeed in the first and advance to the second competition round. It seems that those participants characterized by less developed physics-specific abilities were able to compensate this deficit through their well-developed general cognitive abilities, however, only in the first round. The increased difficulty of the tasks of the second round compared to those of the first round seemed to have had the consequence that compensating a lack of physics-specific abilities with well-developed general cognitive abilities no longer appeared possible. Hence, students who succeeded in the second round and therefore advanced to the third round appeared to be more advanced in their expertise development as indicated through generally more developed physics problem solving abilities (Lind & Friege, 2001; Weinert, 2001).

In sum, it is disputable to what extent the focus of the first competition round aligns to the competition's intention of recognizing and valuing the efforts of engaged and motivated students. The first round appears to identify students characterized by well-developed general or physics-specific cognitive abilities, or both. Engaged and motivated students lacking these cognitive prerequisites have difficulties coping with the competition's demands. The second round seems to identify participants with highly developed problem solving abilities and strong beliefs in their own abilities. Hence, the competition definitively meets its intention of identifying the most capable students, yet it falls short in adequately recognizing and valuing the efforts of engaged and motivated average-ability students.

Implications for improving science competitions

The first round of the Physics Olympiad appears to be overly challenging in the sense that being engaged and motivated alone does not provide participants a reasonable chance to succeed in the entry round of the competition. Hence, to effectively meet the intention of recognizing and valuing the efforts of engaged and motivated students, one may re-evaluate both the difficulty and the types of problems featured in the first round of the competition.

Simply reducing the overall difficulty of the first competition round may result in a greater number of engaged and motivated average-ability students reaching the second competition round, consequently feeling recognized for their efforts. However, it is crucial to ensure that not only more participants advance to the second round but that those who advance do so because of their engagement and motivation. This necessitates a re-evaluation of the types of problems currently employed in the first round. Presently, problems heavily rely on the identification of problem-relevant physics concepts, their mathematical representation, and subsequent mathematical computations to derive a solution. One possible approach could involve replacing a conventional, well-defined problem with a more open-ended one, addressing socially relevant issues intertwined with physics (e.g., "wicked problems", socioscientific issues, see Ramaley, 2014; Zeidler & Nichols, 2009, resp.), since these are the kind of problems that require effort, engagement, and motivation instead of raw physics and mathematics abilities.

The Physics Olympiad and similar science competitions could also offer support programs and learning resources that facilitate the entry into the competition. This way, engaged and motivated students are given an opportunity to learn and practise the knowledge and abilities that are relevant in the first competition round beforehand. Given that physics problem solving abilities consistently emerged as the leading predictor of success in both the first and second round of the Physics Olympiad, we propose the implementation of support programs and learning resources focussing on enhancing this ability prior and during the first and second competition round. Notably, this approach primes students for the competition's demands and also lays a foundation for potential STEM careers in which problem solving abilities are of central importance (Armour-Garb, 2017; Jang, 2016; Mulvey & Pold, 2020). These support programs or resources could explicitly address problem solving strategies (Larkin & Reif, 1979), given their established positive connection to academic achievement (Binder et al., 2019). Moreover, these programs and resources could elaborate on the process inherent to solving domain-specific problems. This could be achieved by presenting learners with a model of the problem solving process (e.g., Polya, 1945; Selçuk & Çalýskan, 2008) encompassing comprehensive instructions for each step of the outlined process. In the domain of physics, this approach has been found to positively influence the quality of students' problem representations (Huffman, 1997; Savelsbergh et al., 1997), their overall problem solving performance and their physics achievement (Selçuk & Çalýskan, 2008).

Lastly, we discuss a more direct approach to recognizing and valuing the efforts of motivated and engaged students. Currently, students participating in the competition engage with the first-round problems over a long period of time alongside their regular school commitments. Ultimately, their physics teachers score their solutions and communicate whether a student advances to the next round based on the achieved score. Advancement to the next round serves as a form of recognition of a student's efforts. However, if a student does not advance to the next round, then the only way of recognizing and valuing this student's efforts is through the teacher who scored the student's solutions. We argue that that it is crucial to make these teachers aware that they hold the key for recognizing and valuing their students' efforts as only they actually see their efforts in the form of the submitted solutions. We therefore propose that teachers offer constructive performance feedback based on their students' submitted solutions to provide recognition and enable future performance improvements (Ellis et al., 2006).

Limitations

We had to refrain from statistical analysis of predictors determining success in the third and fourth competition round of the Physics Olympiad due to too small sample sizes for sound statistical analyses. Hence, we were not able to present any evidence that the trend of identifying the most capable students continues in the third round of the Physics Olympiad. Moreover, our regression analysis did not consider possible interaction effects between variables even though specific combinations of predictor variables might be particularly advantageous for success in the competition. In light of the number of predictors and the larger number of possible interaction terms, considering interactions in a regression framework would probably not yield meaningful results given the sample size. However, using a more holistic approach such as latent profile analysis might unravel interactions between predictors in future analyses (Tschisgale et al., 2024). Lastly, our analysis assumed that predictor variables were stable over the investigated time period. This assumption, however, may not hold as– for example– cognitive variables may change due to learning effects. Future investigations should therefore assess the relevant predictor variables at each competition round for more valid conclusions.

Conclusion and future research

Nowadays, science competitions intend to (1) identify those students with the highest domain-specific cognitive abilities and (2) recognize and value the efforts of engaged and motivated average-ability students. However, our study's findings shed light on a nuanced reality. The first round of the Physics Olympiad seemingly erects a hurdle for students that are engaged and motivated but lack sufficient cognitive abilities, thus challenging the realization of the second intention. Conversely, the Physics Olympiad appears to effectively align with its first intention by successfully identifying students with exceptional physics-specific abilities.

Building on these insights, we advocate for overthinking the kind of tasks employed in the first competition round in order to align the competition with its intention of recognizing and valuing the efforts of engaged and motivated students. Furthermore, we propose the integration of support programs within the competition framework, with a strategic focus on cultivating problem solving abilities. This approach not only readies participants for the competition's demands but also nurtures an ability essential to potential STEM careers. As a call for further investigations, future research should compare the situation to other science competitions beyond the German Physics Olympiads. This comparative approach will enrich our understanding of success in diverse science competitions while also contributing to the generalizability of our findings and leading to a far-reaching evaluation of whether science competitions worldwide meet their stated intentions.

Abbreviations

STEM	science, technology, engineering, and mathematics
M	mean
SD	standard deviation
MAP	maximum a posteriori (estimate)
HPDI	highest posterior density interval
EXSU	expectancy of success
VACO	values assigned to the competition
SEEF	self-efficacy
SOSU	social support
GCAB	general cognitive abilities
PPSA	physics problem solving ability
WLE	weighted likelihood estimate
R1	first competition round
R2	second competition round
R3	third competition round

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s43031-024-00102-y.

Supplementary Material 1

Acknowledgements
Not applicable.

Author contributions
PT wrote the manuscript with substantial input from all other authors and performed statistical analyses. AS contributed significantly to writing the introduction and discussion. MK gave valuable ideas regarding statistical analyses. PW, AS, SP, and KN contributed to the research study design and provided the original data. All authors commented on the manuscript at all stages and helped refining the manuscript across multiple rounds of editing. All authors read and approved the final manuscript. *Language editing.* In the process of refining the manuscript, language editing was conducted with the assistance of ChatGPT, a language model developed by OpenAI. ChatGPT was employed to enhance the clarity, coherence, and style of the text while maintaining the integrity of the original content.

Funding
This work was supported by the Leibniz Association, Germany under Grant K194/2015.

Data availability
The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The data used in this study was collected within the WinnerS project. Participation was voluntary and all ethics requirements for human subjects' research were met as testified by the ethics committee of the IPN under the approval number 2022_13_HO.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 22 November 2023 / Accepted: 21 February 2024
Published online: 04 March 2024

References

Abernathy, T. V., & Vineyard, R. N. (2001). Academic competitions in science: What are the rewards for students? *The Clearing House: A Journal of Educational Strategies Issues and Ideas*, *74*(5), 269–276. https://doi.org/10.1080/00098650109599206.

Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology*, *77*(5), 598–614. https://doi.org/10.1037/0021-9010.77.5.598.

Adams, R. J. (2005). Reliability as a measurement design effect. *Studies in Educational Evaluation*, *31*(2–3), 162–172. https://doi.org/10.1016/j.stueduc.2005.05.008.

Archer, L., Calabrese Barton, A. M., Dawson, E., Godec, S., Mau, A., & Patel, U. (2022). Fun moments or consequential experiences? A model for conceptualising and researching equitable youth outcomes from informal STEM learning. *Cultural Studies of Science Education*, *17*(2), 405–438. https://doi.org/10.1007/s11422-021-10065-5.

Armour-Garb, A. (2017). *Bridging the STEM skills gap: Employer/educator collaboration in New York*. The Public Policy Institute of New York State, Inc.

Avraamidou, L. (2020). Science identity as a landscape of becoming: Rethinking recognition and emotions through an intersectionality lens. *Cultural Studies of Science Education*, *15*(2), 323–345. https://doi.org/10.1007/s11422-019-09954-7.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191.

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

Beauducel, A., & Kersting, M. (2002). Fluid and crystallized intelligence and the Berlin Model of Intelligence structure (BIS). *European Journal of Psychological Assessment*, *18*(2), 97–112. https://doi.org/10.1027/1015-5759.18.2.97.

Binder, T., Sandmann, A., Sures, B., Friege, G., Theyssen, H., & Schmiemann, P. (2019). Assessing prior knowledge types as predictors of academic achievement in the introductory phase of biology and physics study programmes using logistic regression. *International Journal of STEM Education*, *6*(1), 33. https://doi.org/10.1186/s40594-019-0189-9.

Blankenburg, J. S., Höffler, T. N., Peters, H., & Parchmann, I. (2016). The effectiveness of a project day to introduce sixth grade students to science competitions. *Research in Science & Technological Education*, *34*(3), 342–358. https://doi.org/10.1080/02635143.2016.1222361.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Lawrence Erlbaum Associates, Inc.

Boone, W. J., & Noltemeyer, A. (2017). Rasch analysis: A primer for school psychology researchers and practitioners. *Cogent Education*, *4*(1), 1416898. https://doi.org/10.1080/2331186X.2017.1416898.

Brandenburger, M. (2016). *Was beeinflusst den Erfolg Beim Problemlösen in Der Physik? Eine Untersuchung Mit Studierenden [What influences success in problem solving in physics? An investigation with students]*. Logos-.

Bürkner, P. C. (2017). Brms: An R package for bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01.

Campbell, J. R. (1996). Early identification of mathematics talent has long-term positive consequences for career contributions. *International Journal of Educational Research*, *25*(6), 497–522. https://doi.org/10.1016/S0883-0355(97)86728-6.

Campbell, J. R., & Feng, A. X. (2010). Comparing adult productivity of American mathematics, chemistry, and physics olympians with Terman's longitudinal study. *Roeper Review*, *33*(1), 18–25. https://doi.org/10.1080/02783193.2011.530203.

Campbell, J. R., & O'Connor-Petruso, S. A. (2008). *National competitions help eradicate gender inequities in the gifted and talented*. Annual Meeting of the European Council for High Ability, Prague, Czech Republic.

Campbell, J. R., & Walberg, H. J. (2011). Olympiad studies: Competitions provide alternatives to developing talents that serve national interests. *Roeper Review*, *33*(1), 8–17. https://doi.org/10.1080/02783193.2011.530202.

Campbell, J. R., Wagner, H., & Walberg, H. J. (2000). Academic competitions and programs designed to challenge the exceptionally talented. In K. A. Heller, F. J. Mönks, R. Subotnik, & R. Sternberg (Eds.), *International Handbook of Giftedness and Talent* (2nd ed., pp. 523–535).

Chang, C. Y., & Lin, P. L. (2017). The relationship between science achievement and self-concept among gifted students from the third International Earth Science Olympiad. *EURASIA Journal of Mathematics Science and Technology Education*, *13*(7). https://doi.org/10.12973/eurasia.2017.00768a.

Cirik, I. (2015). Relationships between social support, motivation, and science achievement: Structural equation modeling. *Anthropologist*, *20*(1,2), 232–242.

Coleman, E. B., & Shore, B. (1991). Problem-solving processes of high and average performers in physics. *Journal for the Education of the Gifted*, 366–379. https://doi.org/10.1177/016235329101400403.

R Core Team (2021). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/.

Czerniak, C. M. (1996). Predictors of success in a district science fair competition: An exploratory study. *School Science and Mathematics*, *96*(1), 21–27. https://doi.org/10.1111/j.1949-8594.1996.tb10208.x.

DeMars, C. (2010). *Item response theory*. Oxford University Press.

U.S. Department of Education (2013). *For each and every child—A strategy for education, equity and excellence*.

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. https://doi.org/10.1177/0146167295213003.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75–146). W. H. Freeman and Company. http://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/year/2009/docId/12327.

Ellis, S., Mendel, R., & Nir, M. (2006). Learning from successful and failed experience: The moderating role of kind of after-event review. *Journal of Applied Psychology*, *91*(3), 669–680. https://doi.org/10.1037/0021-9010.91.3.669.

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.

Eremin, V. V., & Gladilin, A. K. (2013). International chemistry olympiad and its role in chemical education. *Russian Journal of General Chemistry*, *83*(4), 830–838. https://doi.org/10.1134/S1070363213040373.

Ericsson, K. A. (2018). An introduction to the second edition of the Cambridge handbook of expertise and expert performance: Its development, organization, and content. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (2nd ed., pp. 3–20). Cambridge University Press.

European Commission (2023). *EU Contest for Young Scientists (EUCYS)*. https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/eucys_en.

Ferla, J., Valcke, M., & Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences*, *19*, 499–505. https://doi.org/10.1016/j.lindif.2009.05.004.

Forrester, J. H. (2010). *Competitive science events: Gender, interest, science self-efficacy, and academic major choice* [Doctoral dissertation]. http://adsabs.harvard.edu/abs/2010PhDT.......108F.

Ganotice, F. A., & King, R. B. (2014). Social influences on students' academic engagement and science achievement. *Psychological Studies*, *59*(1), 30–35. https://doi.org/10.1007/s12646-013-0215-9.

Garland, R. (1991). The mid-point on a rating scale: Is it desirable? *Marketing Bulletin*, *2*, 66–70.

Grund, S., Lüdtke, O., & Robitzsch, A. (2016). Pooling ANOVA results from multiply imputed datasets: A simulation study. *Methodology*, *12*(3), 75–88. https://doi.org/10.1027/1614-2241/a000111.

Guo, J., Nagengast, B., Marsh, H. W., Kelava, A., Gaspard, H., Brandt, H., Cambria, J., Flunger, B., Dicke, A. L., Häfner, I., Brisson, B., & Trautwein, U. (2016). Probing the unique contributions of self-concept, task values, and their interactions using multiple value facets and multiple academic outcomes. *AERA Open*, *2*(1), 233285841562688. https://doi.org/10.1177/2332858415626884.

Heller, K. A., & Perleth, C. (2007). *MHBT-S - Münchner hochbegabungstestbatterie für die Sekundarstufe*. [Measurement instrument; MHBT-S - Munich high ability test battery for secondary schools]. Hogrefe.

Höffler, T. N., Köhler, C., & Parchmann, I. (2019). Scientists of the future: An analysis of talented students' interests. *International Journal of STEM Education*, *6*(1), 29. https://doi.org/10.1186/s40594-019-0184-1.

Huffman, D. (1997). Effect of explicit problem solving instruction on high school students' problem-solving performance and conceptual understanding of physics. *Journal of Research in Science Teaching*, *34*(6), 551–570. https://doi.org/10.1002/(SICI)1098-2736)34:6%3C551::AID-TEA2%3E3.0.CO;2-M.

Jang, H. (2016). Identifying 21st century STEM competencies using workplace data. *Journal of Science Education and Technology*, *25*(2), 284–301. https://doi.org/10.1007/s10956-015-9593-1.

Jansen, M., Scherer, R., & Schroeders, U. (2015). Students' self-concept and self-efficacy in the sciences: Differential relations to antecedents and educational outcomes. *Contemporary Educational Psychology*, *41*, 13–24. https://doi.org/10.1016/j.cedpsych.2014.11.002.

Köhler, C. (2017). *Naturwissenschaftliche Wettbewerbe für Schülerinnen und Schüler–Charakterisierung der Anforderungen und Teilnehmenden hinsichtlich spezifischer Leistungsmerkmale* [Doctoral dissertation, University of Kiel; Student science competitions - Characterization of requirements and partcipants with regard to specific performance characteristics]. https://macau.uni-kiel.de/receive/diss_mods_00020955.

Kubsch, M., Stamer, I., Steiner, M., Neumann, K., & Parchmann, I. (2021). Beyond p-values: Using bayesian data analysis in science education research. *Practical Assessment Research and Evaluation*, *26*(1). https://doi.org/10.7275/vzpw-ng13.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159. https://doi.org/10.2307/2529310.

Larkin, J. H., & Reif, F. (1979). Understanding and teaching problem-solving in physics. *European Journal of Science Education*, *1*(2), 191–203. https://doi.org/10.1080/0140528790010208.

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. https://doi.org/10.1119/1.18409.

Lind, G., & Friege, G. (2001). What characterizes participants at the olympiad besides their physics problem solving abilities? Some results from a retrospective survey among former German physics olympiad participants. *Physics Competitions*, *3*(2), 7–15.

Lykkegaard, E., & Ulriksen, L. (2016). Choices and changes: Eccles' expectancy-value model and upper-secondary school students' longitudinal reflections about their choice of a STEM education. *International Journal of Science Education*, *38*(5), 701–724. https://doi.org/10.1080/09500693.2016.1156782.

Madley-Dowd, P., Hughes, R., Tilling, K., & Heron, J. (2019). The proportion of missing data should not be used to guide decisions on multiple imputation. *Journal of Clinical Epidemiology*, *110*, 63–73. https://doi.org/10.1016/j.jclinepi.2019.02.016.

McElreath, R. (2020). *Statistical rethinking: A bayesian course with examples in R and Stan* (2nd ed.). CRC.

Miller, K., Sonnert, G., & Sadler, P. (2018). The influence of students' participation in STEM competitions on their interest in STEM careers. *International Journal of Science Education Part B*, *8*(2), 95–114. https://doi.org/10.1080/21548455.2017.1397298.

Mulvey, P., & Pold, J. (2020). *Physics doctorates: Skills used and satisfaction with employment*. American Institute of Physics.

Neumann, K. (2014). Rasch-Analyse naturwissenschaftsbezogener Leistungstests [Rasch analysis of science-related achievement tests]. In D. Krüger, I. Parchmann, & H. Schecker (Eds.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (pp. 355–369). Springer Spektrum. https://doi.org/10.1007/978-3-642-37827-0.

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology*, *34*(1), 29–48. https://doi.org/10.1080/01443410.2013.797339.

Petersen, S., & Wulff, P. (2017). The German physics Olympiad—identifying and inspiring talents. *European Journal of Physics*, *38*(3), 1–16. https://doi.org/10.1088/1361-6404/aa538f.

Petersen, S., Blankenburg, J. S., & Höffler, T. N. (2017). Challenging gifted students in science—the German Science Olympiads. In K. S. Taber, M. Sumida, & L. McClure (Eds.), *Teaching gifted learners in STEM subjects: Developing talent in science, technology, engineering and mathematics* (pp. 157–170). Taylor & Francis. https://doi.org/10.4324/9781315697147-11.

PISA-Konsortium Deutschland (Ed.). ((2006).). *PISA 2003: Dokumentation Der Erhebungsinstrumente* [PISA 2003 - documentation of survey instruments]. Waxmann.

Polya, G. (1945). *How to solve it—A new aspect of mathematical method*. Princeton University Press.

Ramaley, J. A. (2014). The changing role of higher education: Learning to deal with wicked problems. *Journal of Higher Education Outreach and Engagement*, *18*(3), 7.

Resch, C. (2013). Olympiads as a means to promote gifted students. *Physics Competitions*, *15*(1 & 2).

Robitzsch, A., Kiefer, T., & Wu, M. (2021). *TAM: Test analysis modules* (R package version 3.7–16) [Computer software]. https://CRAN.R-project.org/package=TAM.

Savelsbergh, E. R., Ferguson-Hessler, M. G. M., & de Jong, T. (1997). *The importance of an enhanced problem representation: On the role of elaborations in physics problem solving*. University of Twente Faculty of Educational Science and Technology, Department of Instructional Technology. Instructional Technology Memorandum Series.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*(1), 162–173. https://doi.org/10.1037/0022-3514.86.1.162.

Science Olympiad Inc (2023). *Science Olympiad*. https://www.soinc.org/about/mission.

Selçuk, G. S., & Çalýskan, S. (2008). The effects of problem solving instruction on physics achievement, problem solving performance and strategy use. *Latin-American Journal of Physics Education*, *2*(3), 151–166. http://www.lajpe.org/sep08/01_Gamze_Sezgin.pdf.

Shechter, O. G., Durik, A. M., Miyamoto, Y., & Harackiewicz, J. M. (2011). The role of utility value in achievement behavior: The importance of culture. *Personality and Social Psychology Bulletin*, *37*(3), 303–317. https://doi.org/10.1177/0146167210396380

Simpkins, S. D., Price, C. D., & Garcia, K. (2015). Parental support and high school students' motivation in biology, chemistry, and physics: Understanding differences among latino and caucasian boys and girls. *Journal of Research in Science Teaching*, *52*(10), 1386–1407. https://doi.org/10.1002/tea.21246.

Smale-Jacobse, A. E., Meijer, A., Helms-Lorenz, M., & Maulana, R. (2019). Differentiated instruction in secondary education: A systematic review of research evidence. *Frontiers in Psychology*, *10*, 2366. https://doi.org/10.3389/fpsyg.2019.02366.

Smith, K. N., Jaeger, A. J., & Thomas, D. (2021). Science Olympiad is why I'm here: The influence of an early STEM program on college and major choice. *Research in Science Education*, *51*(S1), 443–459. https://doi.org/10.1007/s11165-019-09897-7.

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 [Who will advance? Predicting qualification for the International Biology and Chemistry Olympiads based on Eccles' model of achievement motivation]. *Zeitschrift Für Pädagogische Psychologie*, *28*(3), 105–114. https://doi.org/10.1024/1010-0652/a000127.

Steegh, A., Höffler, T., Höft, L., & Parchmann, I. (2021). Exploring science competition participants' expectancy-value perceptions and identification: A latent profile analysis. *Learning and Instruction*, *74*, https://doi.org/10.1016/j.learninstruc.2021.101455.

Stoltzfus, J. C. (2011). Logistic regression: A brief primer. *Academic Emergency Medicine*, *18*(10), 1099–1104. https://doi.org/10.1111/j.1553-2712.2011.01185.x.

Tirri, K. (2010). Motivation and giftedness. *High Ability Studies*, *21*(2), 77–80. https://doi.org/10.1080/13598139.2010.528923.

Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Journal of Educational Psychology*, *104*(3), 763–777. https://doi.org/10.1037/a0027470.

Treiber, E., Neumann, I., & Heinze, A. (2023). What's mathematics doing here? The role of mathematics in German physics olympiad tasks. *Frontiers in Education*, *8*, 1196189. https://doi.org/10.3389/feduc.2023.1196189.

Tschisgale, P., Steegh, A., Kubsch, M., Petersen, S., & Neumann, K. (2024). Towards a more individualised support of science competition participants– identification and examination of participant profiles based on cognitive and affective characteristics. *International Journal of Science Education*, 1–25. https://doi.org/10.1080/09500693.2023.2300147.

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. https://doi.org/10.1080/13598139.2012.738324.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3). https://doi.org/10.18637/jss.v045.i03.

van Ginkel, J. R., Linting, M., Rippe, R. C. A., & van der Voort, A. (2020). Rebutting existing misconceptions about multiple imputation as a method for handling missing data. *Journal of Personality Assessment*, *102*(3), 297–308. https://doi.org/10.1080/00223891.2018.1530680.

Verna, M. A., & Feng, A. X. (2002). American chemistry olympians achieve the highest level of equity. *Journal of Research in Education*, *12*, 101–105.

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*(4), 817. https://doi.org/10.1037/a0016127.

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

Warrens, M. J. (2012). Cohen's linearly weighted kappa is a weighted average. *Advances in Data Analysis and Classification*, *6*(1), 67–79. https://doi.org/10.1007/s11634-011-0094-7.

Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen—Eine Umstrittene Selbstverständlichkeit. [Comparative performance measurement in schools - a controversial matter of fact]. In F. E. Weinert (Ed.), *Leistungsmessungen in Schulen* (2nd ed., pp. 17–31). Beltz.

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review*, *12*(3), 265–310. https://doi.org/10.1016/0273-2297(92)90011-P

Wille, E., Stoll, G., Gfrörer, T., Cambria, J., Nagengast, B., & Trautwein, U. (2020). It takes two: Expectancy-value constructs and vocational interests jointly predict STEM major choices. *Contemporary Educational Psychology*, *61*, 101858. https://doi.org/10.1016/j.cedpsych.2020.101858

Wulff, P., Hazari, Z., Petersen, S., & Neumann, K. (2018). Engaging young women in physics: An intervention to support young women's physics identity development. *Physical Review Physics Education Research*, *14*(2), 020113. https://doi.org/10.1103/PhysRevPhysEducRes.14.020113.

Wulff, P., Tschisgale, P., & Petersen, S. (2023). *Physics problem solving test and coding rubric* [Measurement instrument]. https://doi.org/10.17605/OSF.IO/GSMVR.

Zeidler, D. L., & Nichols, B. H. (2009). Socioscientific issues: Theory and practice. *Journal of Elementary Science Education*, *21*(2), 49–58. https://doi.org/10.1007/BF03173684.

Zhou, X., & Reiter, J. P. (2010). A note on bayesian inference after multiple imputation. *The American Statistician*, *64*(2), 159–163. https://doi.org/10.1198/tast.2010.09109.

Publisher's Note