

Engaging young women in physics: An intervention to support young women's physics identity development

Peter Wulff,^{1,*} Zahra Hazari,² Stefan Petersen,¹ and Knut Neumann¹

¹*Leibniz-Institute for Science and Mathematics Education (IPN),
Olshausenstraße 62, 24118 Kiel, Germany*

²*STEM Transformation Institute, Florida International University,
ZEB 259A 11200 SW 8th St., Miami, Florida 33199, USA*

 (Received 9 June 2017; revised manuscript received 3 June 2018; published 5 November 2018)

This study presents findings on the physics identity development of female students in the German Physics Olympiad who participated in an intervention designed to support their engagement in physics. Enrichment programs such as the Physics Olympiad have been found to positively impact students' engagement and intent to pursue a career in science. However, many enrichment programs, including the Physics Olympiad, suffer from an underrepresentation of young women. The intervention investigated in this study capitalizes on gender-related research in physics education in order to explore ways in which gender equity can be raised in enrichment programs. To this end, we designed an identity-safe learning environment that facilitates participating young women's physics identity development. For example, same-sex groupings and considerately adopted physics contents that particularly acknowledge young women's interests (e.g., relation to medical issues and the human body) were utilized. Overall, 30 Olympians took part in a one-day intervention (13 females, 17 males). Positive effects in two important physics identity constructs, namely, *interest* and *competence*, for young women were found, while at the same time no effects were found for young men. Furthermore, the young women were more likely to participate in next year's Physics Olympiad, compared to the overall female Physics Olympiad population. These results indicate that the careful design of an intervention based on gender research and science identity theory can support young women's physics identity development.

DOI: [10.1103/PhysRevPhysEducRes.14.020113](https://doi.org/10.1103/PhysRevPhysEducRes.14.020113)

I. MOTIVATION

Modern societies are driven by ever-accelerating scientific and technological progress [1]. In order to ensure future wealth, health, and sustainability in a globalized world, modern societies must strive for excellence in the fields of science, technology, engineering, and mathematics (STEM) [2]. Recruiting and retaining students to STEM is a major goal for modern societies [2,3]. A particular concern, from this perspective, is that engagement in many STEM domains is largely patterned by gender. Despite decades of research on this problem, many industrialized countries such as Germany and the U.S. show only little growth in the representation of women in STEM [4]. Young women are, for example, still underrepresented in many STEM domains and have disproportionately higher attrition from more influential positions [5]. The underrepresentation and

attrition of young women from STEM domains point to mechanisms of gender inequity in STEM that hamper emancipation of women in technology reliant societies and impair advancement of STEM domains since potential remains untapped [6].

In order to attract young students to STEM and promote their engagement, in addition to formal learning environments, many countries implement informal science learning programs such as science enrichment programs [7,8]. Research suggests that these science enrichment programs can support participating students in the development of competence beliefs [9]. Also, participants retrospectively attribute positive experiences in these programs to driving their motivation for a science career [10]. Examples for such programs are the science Olympiads in which high school students from various countries compete once a year across different science disciplines. In most of these contests, students compete in multiple rounds against each other, first through remote problem solving and eventually through in-person live contests [8]. Students participating in these competitions report positive effects on their interest and career decisions [11]. Furthermore, research on former participants suggests that the vast majority of the

*peter.wulff@uni-potsdam.de

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

competitors in the higher rounds pursue a math or science-related career and exhibit performance above average [12].

However, these programs suffer from underrepresentation and attrition of young women. Especially for the Physics Olympiad, the underrepresentation of young women is well documented. Young women, compared to young men, are fewer in the initial rounds in most countries and have higher rates of drop out from the competition at earlier stages [13]. For example, in many countries typically around 25% of the participants in the primary stages are young women. In the international competition, after the teams have been chosen, only around 7% of the participants are young women [8,14]. Focusing on the target country in this study, gender patterns are a particular concern in the German selection process for the Physics Olympiad compared to many other countries [13,15]. Around 18% to 27% of the participants in the first stage in the German Physics Olympiad are young women, while in the final stage only around 5% are young women [8].

In the past, educators sought to change the young women in their attitudes towards physics (e.g., through re-attributional training [16]). More recent research, however, documents how the social context (e.g., gender representation [17]) threatens young women in their identity development. For example, stereotypes might exacerbate feelings of belongingness for young women or overrepresentation of young men in physics learning environments might induce feelings of mismatch with the context. This may establish barriers to the engagement for young women in physics [18]. Informal learning programs such as the Physics Olympiad place young women in particularly threatening contexts because men outnumber young women by large proportions and no viable strategies with diverse perspectives are set in place to challenge the underrepresentation of young women in physics [8]. In order to ameliorate this situation, it is important to take diverse perspectives into mind when designing interventions such that the programs themselves assimilate diverse goals and values that are more equitable at their core, which eventually encourages more young women to express their identity within these programs.

This study seeks to explore ways to alter the social context of the Physics Olympiad towards a greater alignment with the motivations and goals of young women. As a means to do so we present an intervention with the goal of establishing an identity-safe physics environment for physics-interested young women in the context of the Physics Olympiad in order to support their physics engagement and their physics identity development. Identity-safe physics environment refers to a context where the female gender identity is neither salient nor devalued on the basis of group constellation or stereotypes. The intervention functions as a proof of concept in order to inform future interventions aimed at raising gender diversity in the Physics Olympiad and potentially other physics learning

environments. We chose to implement our study in the context of the Physics Olympiad because the underrepresentation of women is well documented in these spaces [19]. Moreover, developing intervention strategies with physics-interested students is a more practical starting place for changing social contexts to further support women who are somewhat interested in the content already. Engaging students who have no prior experience or interest requires a different informal science education context—contexts which equally deserve attention but are not the focus in this study.

II. THEORETICAL BACKGROUND

Gender issues in math-intensive domains such as physics have been researched from different perspectives [20]. On the one hand, research in the tradition of sociological perspectives embraces rather holistic concepts, like agency and broadly understood identity (in the sense of social roles), to address the problem of underrepresentation. For example, the influence of constraining social structures is prominently discussed in the ways that certain identities and social roles are culturally produced and there is a limitation of agency, understood as the capacity to act intentionally [21], in physics [22]. These approaches emphasize that learning is intricately linked to societal structures such that the sociocultural context plays an important role for understanding individual learning rather than mere cognitive abilities of the individual [23]. On the other hand, research in the tradition of social psychology capitalizes on individual variables, like stereotype endorsement, interest, motivation, or sense of belonging, in order to explain the underrepresentation of young women in math-intensive domains like physics. Here, individual cognition and individual learning are target constructs [16,24]. For example, researchers came to understand the detrimental effects on learning when the students' gender is salient to them in a physics classroom setting or when gender stereotypes are present [25,26]. Both perspectives contribute to our understanding of female underrepresentation in physics. A theory that bridges the individual and environmental (contextual) focus is science identity. Science identity captures important aspects, such as a more holistic understanding of individual development and relates to the underrepresentation of females in STEM [27,28]. In particular, research with an identity lens has been considered to be a “missing link” between learning and its sociocultural context [29].

A. Physics identity

In this work, identity development is defined as the process of searching for and settling on a set of commitments to personal standards and life roles in different social contexts [30]. This process entails the negotiation of multiple identities that a student holds, e.g., social identities

such as a gender identity, personal identities such as the association of personality traits with the self, and disciplinary identities such as the affiliation with a school subject for a student [31]. A student typically acts in alignment with his or her identities in a given setting [18] and ends up being recognized by others as a certain “kind of person” in a context [32]. The focus of this research is on specific disciplinary identity, namely, physics identity. In the process of developing a physics identity the student comes to see herself or himself as a physics person and is seen by others as a physics person [33].

The constructs that are central to a physics identity have been qualitatively established [27] and replicated in subsequent studies [28,34]. The following constructs were found to be integral to physics identity formation for students: *interest*, *recognition*, *competence*, and *performance* (see Fig. 1). *Interest* is the enjoyment the student has in dealing with physics. Early interest in physics in middle and high school was found to be a strong predictor of later academic choices [35] and particularly dependent upon stimulating environments [36]. For math interest it was shown that it was related to taking up of advanced courses in math [37]. An explanatory link for the relationship between interest and persistence has been found to be a student’s identity in a domain, i.e., when students become more interested, they start to see themselves as that “kind of person,” and ultimately choose to persist. Recognition is the students’ perception of how much others see her or him as a physics person. Recognition is particularly linked to social cues. For example, the recognition by meaningful others (parents, peers, and teachers) as a “physics person” is strongly related to having a positive perception of the domain [38]. Recognition by others as a “physics person” in high school has been established to correlate with

physics identity and intended physics career [28]. Furthermore, the lack of recognition can lead to disrupted identities [27]. Competence entails the students’ belief in their ability to be good at the required tasks and understand physics. Competence has been established to be important at the outset of engagement in a domain [39]. Self-efficacy in a domain, which is similar to competence beliefs, although more task specific, was found to be predictive of students’ performance and later educational outcomes, and vice versa [40]. The feeling of competence can also vary in relation to what social context a student faces (e.g., social comparisons). In particular, in gifted and talented programs the perceived competence can drop [41]. Performance refers to educational outcomes, where it was shown that good performance is a prerequisite to build a positive self-concept [42]. When performance is operationalized as a performance belief (i.e., the belief to be able to perform well in physics), it was shown that competence and performance form one construct for students [43].

In sum, interest, recognition, and competence or performance (from here on competence) facilitate students’ identity development in physics. Integrating these constructs into one framework (Fig. 1) enables researchers to better understand students’ academic participation and persistence in physics, and its relation to the social context (e.g., the Physics Olympiad) and students’ experiences in learning environments (e.g., interventions). It has been shown that this identity configuration holds empirical strength in explaining and predicting academic choices in STEM [28,34].

B. Gender and physics identity

When considering gender-related patterns in physics enrollment, on average, boys and young men, compared to girls and young women, more readily identify with

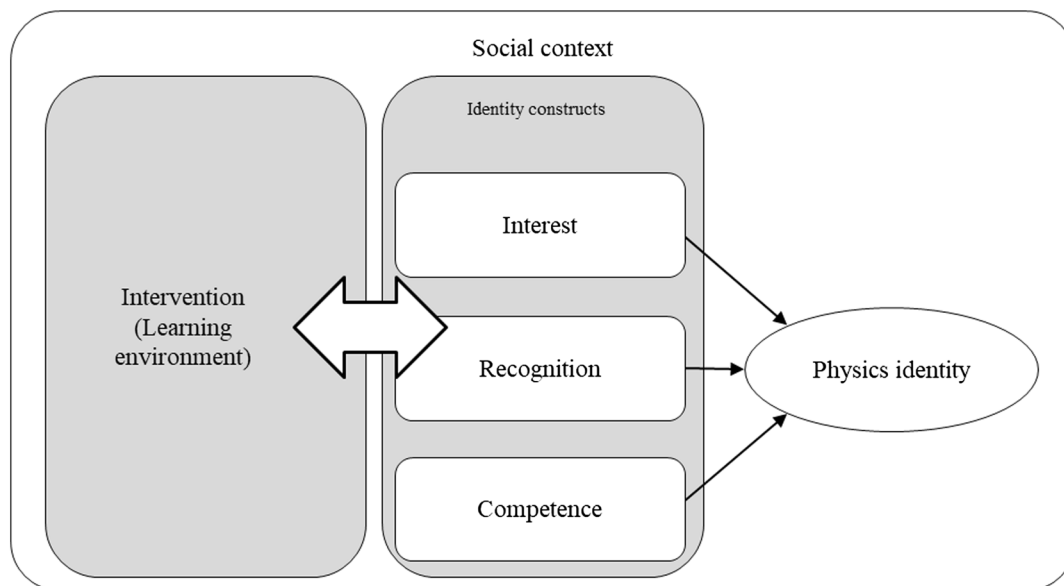


FIG. 1. Physics Identity Framework (adapted from Refs. [28,44]).

math-intensive domains like physics and place a higher personal value on these domains [45]. The reasons for this are explained, at least in part, when considering the identity constructs of interest, recognition, and competence. Boys have been found to express an interest to deal with physics stuff, in both school and leisure time, more than girls [46,47]. Girls have been found to lose their interest in physics at early ages [46,48]. Some of the disengagement is accounted for by early childhood play. For example, boys have more exposure and experience with toys and objects that align with the traditional physics curriculum than girls do [49,50]. In particular, a lack of real-world connections and personally relevant content have been found to result in young women disliking physics [28,49].

Considering recognition, even successful women in science were found to lack support in their surroundings that would lend to feelings of recognition [27]. Brickhouse and Potter highlight that girls in physics classrooms often express a feeling of isolation and experience a chilly climate [51]. Similarly, other studies have found that physics classrooms are often dominated by boys [52] making it more difficult for girls to perform and be recognized for what they know and can do. Furthermore, the female gender identity is particularly threatened when it comes to performance situations in physics [25]. Persistent social stereotypes frame math-intensive domains like physics as something for (male) geniuses and brilliant people [53]. Such stereotypes about women not being as competent in certain domains, such as physics, can impair performance [54]. Physics environments can even be seen as “hazardous” by the young women who experience them [55]. Such conflicts with ones’ gender identity eventually lead to disengagement with the domain [24,56]. Gender stereotypes in physics are particularly problematic for young women [57], because they limit what female students believe they can do in the field of physics, in addition to their disinterest when their goals do not match physics-related careers to begin with [58].

Abundant gender-linked research exists specifically relating to competence beliefs. With regards to achievement, which is correlated with competence, male students in college tend to outperform females on conceptual understanding assessments [59]. Yet, this is not due to innate differences in cognitive ability [20,60]. Prior research has found that boys have an advantage in physics experiences outside of school and that these experiences can account for perceptions of higher competence as well as higher performance in physics [61]. Young women hesitate to fully engage with physics. They tend to perceive physics knowledge in schools as heteronomous and more difficult than other subjects [62], as well as objective and valid for all time [63]. As indicated earlier, ability stereotypes for young women in physics learning settings can hamper their performance [25,64]. Consequently, in the course of their schooling, young women, compared to

young men, acquire depressed beliefs in their ability to perform well in physics [65].

Many researchers have sought to intervene and empower young women to pursue physics. Carefully designed, evidence-based intervention studies that consider the physics identity constructs were able to produce positive results with regards to interest and engagement for young women. For example, Häußler and Hoffmann showed that the adaptation of the curriculum in physics classrooms to the specific interests of girls had positive effects on physics self-concept, a construct related to physics identity (competence) [66]. Berger adopted a novel physics curriculum with contextualized materials (x ray in medicine) and compared students’ interest and performance to a traditional curriculum [67]. Particularly the young women who lacked an initial interest were found to benefit the most from the new curriculum in terms of interest and performance gains. Also, an active physics classroom (e.g., through hands-on experiments) had positive effects on engagement [68]. Small-scale interventions can also have positive effects for students that identify with groups that are historically marginalized in physics learning settings [69]. Such small-scale interventions account for the potentially threatening nature of certain situational cues in social learning settings by providing strategies to the students that better protect the targeted students from identity threats. An important message from these interventions is that even small-scale interventions that target processes of identity threat can show abiding effects even after years for students that identify with marginalized groups [70]. An integrated approach that brings together approaches from multiple interventions to facilitate engagement for young women in physics settings and promote their physics identity development may be most beneficial, and thus provide strategies for the physics community to tackle the problem of female underrepresentation in a more holistic and meaningful way.

III. RESEARCH QUESTIONS

In order to explore ways for altering the social context of the Physics Olympiad towards more gender equity and thus facilitating physics identity development for young women through such informal science programs, we designed a research-based intervention in the context of the selection process of the Physics Olympiad in Germany. This intervention is meant to facilitate physics identity development for participants of all sexes in the Physics Olympiad with a particular focus on young women in the competition. In order to do so, the intervention strives to provide an identity-safe context for the young women who already participate in the Physics Olympiad where they experience themselves as competent problem solvers. Overall, the students were given a one day in-person intervention where they came together and worked on physics problems in a carefully designed learning environment. The intervention’s target population

is adolescent students, in particular young women who participated in the Physics Olympiad. Researchers have argued that high school, which is the time period in which the Olympiads occur, is an important moment for engaging young women with physics and tackling the problem of underrepresentation [71]. These young students are in their adolescence and thus move through a developmental period where they discover and develop their (gendered) identities [72,73]. In this developmental stage, adolescents consolidate their commitments to life plans and projects [73,74], and young women build their motivation to become physicists [75]. Besides the academic strength of the young women in the intervention and their potential to excel in physics, these students tend to be affected by messages from the social environment (e.g., ability stereotypes) [76]. It has been argued that young women with an academic strength in science or mathematics might be especially responsive to STEM-related interventions [77] with regards to their identity. The physics identity constructs informed the design for the intervention and the research questions are tailored to the physics identity constructs and to indicators of engagement in the Physics Olympiad.

RQ 1: How does the physics identity of young women and men who participate in the intervention develop where the intervention was designed to specifically address the physics identity constructs interest, recognition, and competence?

Besides the measures for identity constructs, we were interested whether the students changed their attitudes towards the Physics Olympiad itself since this would indicate important effects of the intervention. In other words, the intervention will be considered a proxy for the Physics Olympiad, such that a positive experience in the intervention might translate into a more positive (negative) attitude towards further participation in the Physics Olympiad itself. Ultimately, we also tracked whether the students enrolled in the next year's Physics Olympiad as an indicator for the students' affiliation with the Physics Olympiad.

RQ 2: To what extent do participating young women and men change their participation (intended and actual participation) in the Physics Olympiad?

IV. METHODOLOGY

A. Setting

Approximately 1000 students from more than 300 schools in Germany participate each year in the German Physics Olympiad in four subsequent rounds that tax, amongst others, the students' physics problem solving skills. In this process, we implemented the one-day intervention between the first and second round of the 2015's Physics Olympiad, because at this point the participants are most amenable regarding their potential future engagement in the Physics Olympiad. For example, only approximately half the students who are young enough to participate in the next year's Physics Olympiad do so.

Three different locations (high schools and research sites) were chosen where the students received the intervention in three separate groups, each at one of the locations in November 2015. The intervention took approximately six hours on a single day. First, a small introductory game to get to know each other was played and same-sex groups were formed. These groups were seated at group tables and a discussion session was available for phases in which the whole group discussions took place (e.g., when results to the problems were compared). The students then worked through topics in radiation physics as an essential topic in the curriculum of the Physics Olympiad: waves and oscillations, light as electromagnetic wave, properties of waves, and an experiment on light diffraction. In order to test the effects for the physics identity constructs (RQ 1), the participants took pre- and postsurveys with all relevant measures before and after the intervention, as will be detailed later. The presurvey took approximately 50 min and the post survey took approximately 30 min, additional to the 6 h of the intervention time. The surveys were administered right before and after the instruction took place. We furthermore tracked the students' participation in the next year's Physics Olympiad that started in April 2016. Typically, around 40%–45% of the students who are young enough in the population of the Physics Olympiad also participate in the next year's Physics Olympiads.

B. Design of the Intervention

Steele noted that it is important to “render[] onto the right students the right interventions” [76]. In order to design the right intervention we reflected upon the findings of prior research. We identified studies with findings that relate to the physics identity constructs interest, recognition, and competence. The integration of appealing and relevant physics content, a careful consideration of the social environment (e.g., constellation of groups), and the instructional strategies are each closely linked to the identity constructs. These different design aspects of the intervention will be considered in detail in the remainder of this section. See Fig. 2 for an overview of the particular design elements that were adopted in the intervention. We do not assume single paths from the design elements to certain constructs of physics identity. For example, in referencing self-determination theory, Krapp points to the interrelatedness of competence beliefs in a setting and the resulting interest to deal with content in this domain [78]. Similarly, same-sex groupings during an intervention might affect recognition by the community and competence beliefs at the same time.

Motivating and relevant physics content.—Young women and men differ in their motives to pursue physics and in their particular interests in different physics content. For example, Seymour and Hewitt show that a consistent motive for young women to switch STEM subjects away from physics was that they want to help other people [79].

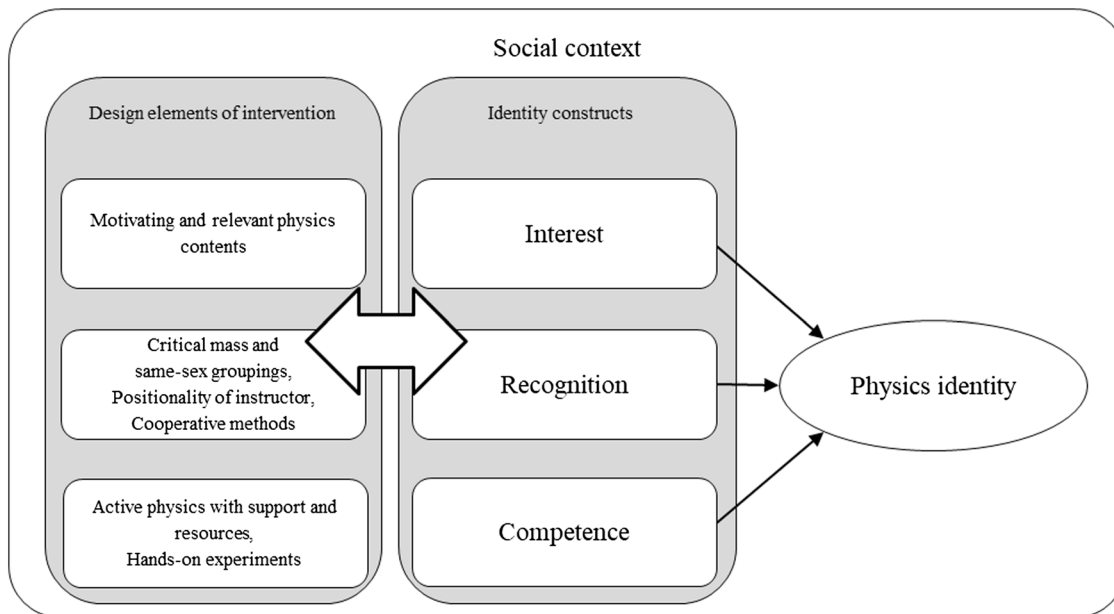


FIG. 2. Design elements of the intervention.

Young women preferably choose biological science and medicine from amongst the STEM fields [80]. These insights into occupational motives are well replicated with interest studies in physics. Hoffmann *et al.* [49] and the ROSE study across multiple countries [81] are just two examples. Young women particularly show an interest in medical-biological contexts such as “cancer—and how we can treat it” [82]. Stadler *et al.* remind us that many contexts that are particularly interesting to females are also interesting to male students [83]. For example, young men have been found to rate medicine as equally high in terms of interest compared to technical contexts [84]. For this intervention we focused on structural analysis using diffraction and particularly tied in the context of the use of x-ray diffraction in uncovering the structure of human DNA. This emphasizes the social relevance of physics knowledge and it is a context that is directly linked to biology, medicine, and the human body. Rosalind Franklin was introduced to the students as the scientist who paved the way to the structural analysis of human DNA with her famous diffraction picture of DNA (“Photo 51”). In order to realize an experimental setup that the students could safely use, rather than x ray, we used visible light. Our content was guided by research and the International Physics Olympiad 2015 problems that utilized an experimental problem related to this content [85]. For example, a laser-pointer was used to model the radiation and thin gauge wire for the structure of DNA. As such, this experiment directly relates to important life-science aspects that are part of physics as well.

Critical mass and same-sex groupings: In order to facilitate more opportunities for *recognition*, same-sex small group work was utilized since men in mixed-gender groups have been found to dominate and the intellectual

contributions of female students suffer in these constellations [26,86]. Having same-sex groupings was proposed to allow women in the small group work to be able to recognize each other and ensure that they had a voice in the activities [17,86]. The small groups ultimately came together to collectively share ideas and findings with the whole seminar group providing further opportunities for being recognized. Furthermore, research suggests that the male-to-female ratio is an important feature to limit identity threats and facilitate recognition [86]. The male-to-female ratio in groups has been found to make one’s gender salient which leads to more gender-stereotypical behavior, i.e., men dominating the discussion with women receiving less recognition for their ideas [26]. When gender becomes salient for young women in physics settings their performance is negatively affected [25]. Young women in same-sex classes, for example, were also shown to have a greater interest in science than those in mixed-sex classes [87]. In order to lower the salience of the students’ gender identity, we sought to bring together young women and men in equal ratios (50:50) for the entire intervention, although our sample was slightly unbalanced (13 female, 16 male). Ratios of 50:50 have been found to be as effective as ratios in which young women were in the majority [17]. Having 50:50 ratios instead of a single-sex intervention also enabled us to make the environment a little more authentic to what students would experience in terms of the male gender representation in the Olympiad (e.g., in the later rounds of the Olympiad, female representation is usually less than 15%). The intervention group consisted of enough young women to form more of a critical mass, i.e., a sufficient representation such that a student’s gender does not become salient and impair her engagement.

Cooperative methods.—Potvin and Hasni, in their review of empirical research on cooperative learning (i.e., “learners interacting and working together to facilitate acquisition or problem solving, by sometimes sharing experiences and knowledge”, p. 104), established that cooperative learning has a positive effect on interest, motivation, and attitude [88]. In particular, female students have been found to appreciate cooperative learning methods [89]. For example, especially for young women in early adolescence, contacts with friends occur more frequently compared to young men [90] and young women, on average, have a strong inclination towards social aspects, e.g., communication [91]. Also, women tend to place a stronger focus on interpersonal relations in their developing identities [92]. In the intervention, the students were encouraged through instruction to work together in pairs and with four students interchangeably.

Positionality of instructor.—Reflecting on positionality is a crucial aspect when considering engagement for marginalized groups in physics. According to Davies and Harré, positionality accounts for the allocation of responsibility and status in local social settings [93]. While the instructor in our intervention might be seen as a prototypical representative for the competition context (white male Ph.D. student), he was also reflective of his positioning and conscious of relevant issues that were found to depress young women’s physics identity development. Drury *et al.* encourage the idea that males should function as well as role models and mentors so that diversification is not seen as a female issue but rather a societal issue [94]. As such, a representative of the competition context can also be seen as a mentor [53,95].

Active physics and hands-on-experiments.—In the intervention the physics concepts that were dealt with (focusing on light diffraction) were divided into coherent learning sequences. The students were provided materials and experiments in order to give them the opportunity to meaningfully engage with learning the concepts. Phases of cooperative work were usually followed by a coming together of students in bigger groups in order to share the results. It was emphasized for students to draw their own conclusions and generate insight (active physics classroom [96]). As suggested in active physics instruction, qualitative understanding of concepts was given considerable attention throughout the intervention [97]. The role of the instructor was to support the students in their learning process [66]. For women, compared to men, a known challenge is the often documented difference in prior experiences and background knowledge in physics [48,98]. This lack of prior experience with physics seems to be one reason that young women particularly benefit from carefully constructed active learning scenarios and hands-on experiences [99]. This is because active learning can particularly reduce the reliance on prior-knowledge and enable the learners with first-hand experiences in the course of the learning

activity. Thus, all participants were provided with opportunities to experience and engage with the physics phenomena as well as share their results. For those who were quick to understand the concepts and carry out experiments, additional activities were provided.

V. INSTRUMENTS

In order to test the effectiveness of the intervention, a host of constructs was surveyed. In order to do so instruments were utilized from science education, in particular science identity research. These instruments were adapted to the Physics Olympiad context, since, to our knowledge, no prior interventions in informal science settings have been designed to analyze the development of the identity constructs. Table I gives an overview of the constructs surveyed together with additional information such as the time the instrument was administered, the number of items (No.), the internal consistency as measured through Cronbach’s α (or Spearman-Brown for 2-item scales), and one sample item.

A. Interest

Interest is not a monolithic construct [100], but rather conceptualized on the basis of different aspects that might be concealed in aggregated scales [49]. As indicated in Table I the different interest scales are *Content interest in physics*, *Physics class interest*, and *Situational interest (post)*. A general physics interest (scale: Content interest in physics) was used in order to measure baseline physics interest [101]. Since we did not expect a single intervention to change the students’ content interest in physics, this was measured only prior to the intervention as a baseline comparison for participants. Such broad measures change in the course of schooling, but not due to one-time interventions [78]. The corresponding items were measured on a 4-point Likert scale (1: “disagree” to 4: “agree”). The internal consistencies are satisfactory for these scales (see Table I). Finally, we measured interest for physics class (scale: Physics class interest) with an item where the students rated how interesting they find their physics class. The item is displayed in Table I. The responses for this item were measured on a 5-point Likert scale (1: “do not like it” to 5: “enjoy it very much”). The interest for school physics classes is a confounding variable and is included to rule out group differences between students with respect to their interest for their physics class. It has been argued that content interest and course interest in physics are distinct from each other [49]. In order to check how interested the students were in the intervention topic, a scale by Fechner [102] that measures situational topic related interest was used as a post measure and we called it Situational interest (post). See a sample item in Table I (as for all other scales). The items were measured on a 5-point Likert scale (1: “not true” to 5: “true”).

TABLE I. Overview of instruments that were measured in the intervention.

Construct	Scale	Time	No.	α	Sample item
Identity constructs:					
Interest	<i>Content interest physics</i>	1	5	0.84	“In general I enjoy doing physics.”
	<i>Physics class interest</i>	1	1		“How much do you enjoy your physics classes in school.”
	<i>Situational interest (post)</i>	2	6	0.83	“The topic of the seminar was very interesting to me.”
Recognition	<i>Recognition</i> in Physics Olympiad	1,2	2	0.45, 0.59	“I feel that the IPhO-team recognizes my engagement in the Physics Olympiad.”
	<i>Recognition</i> in physics class	1,2	2	0.59, 0.60	“I feel that my physics teacher recognizes my engagement in the physics class.”
Competence	<i>Competence</i> belief in Physics Olympiad	1,2	2	0.80, 0.96	“I feel competent to solve the tasks in the Physics Olympiad.”
	<i>Competence</i> belief in physics class	1,2	1		“I feel competent to solve the tasks in the physics class.”
Further engagement:	Intent of further participation	1,2	1		“How likely will you participate in the next year’s Physics Olympiad?”
	Actual participation in next year’s Physics Olympiad	3			
Covariates	Self-reported gender	1	1		
	School Subjects (grades and interest)	1			“How much do you like [subject]?”
	Parental background	2			“What is the highest educational degree of your mother?”
	Support by teachers, parents, and peers	1	4	0.66, 0.69, 0.69	“My teacher (parents, peers) actively support(s) me in my physics engagement”
	Performance in Physics Olympiad	1	4		
	Content knowledge (CK)	1,2	14		

B. Recognition

For recognition we came up with separate measures for the contexts of the Physics Olympiad and physics class, since recognition in the one context might differ from the other context. Recognition was measured pre and post on a 4-point Likert scale (1: “untrue” to 4: “true”). The physics class recognition items were similar to the items for the Physics Olympiad, just replacing Physics Olympiad with physics class, and Physics Olympiad team with physics teacher. The Spearman-Brown formula was used to calculate the internal reliability, since it more appropriately measures two-item reliability than Cronbach’s α [103]. The reliabilities for recognition are rather unsatisfactory (see Table I) as will be discussed later.

C. Competence

As with recognition, the competence beliefs measures were tailored to the competition context (Table I) since we did not expect the students to develop broader physics competence beliefs as a result of a single intervention. It has been argued elsewhere that scales such as self-concept (which is similar to competence) should be conceptualized with regards to the context [104]. Furthermore, it is different for a student to feel competent in a physics

competition context as compared to a physics classroom context, since the Olympiad context comprises physics-interested students that eventually compete with each other [104]. We adopted items that closely fit the descriptions of the constructs within the science identity literature [28] (see Table I). Competence was measured pre and post on a 4-point Likert scale (1: “untrue” to 4: “true”). The reliabilities are satisfactory. Additionally, we measured competence beliefs for physics class. Only one item addressed competence in physics class. One item scales are comparable to multiple item scales when the construct is simple, and has a single-meaning attribute (e.g., liking) [105]. Usually competence items fulfill these requirements and show very high internal reliability [105]. The item reads “I feel competent to solve problems in my physics class.” It is measured on a 4-point Likert scale (1: “untrue” to 4: “true”).

D. Engagement in Physics Olympiad

To measure intentions to persist in the Physics Olympiad, students placed a cross on a continuous scale that indicated how likely they thought they were to participate in the next-year’s Physics Olympiad. The anchors on the scale were “not likely” and “very likely.”

The item was scored based on the distance from the “not likely” anchor divided by the total length of the scale. In order to account for the measurement uncertainty of the subjective choice the responses were binned. In order to do so the scale was cut into 8 equidistant intervals and each response was classified accordingly based on its measurement for its position on the line. Students’ intentions to persist were measured pre and post. Furthermore, in order to see whether students further engaged in the competition, we tracked the students’ enrollment in the next year’s Physics Olympiad.

E. Covariates

Several background variables were used in order to ascertain comparability of male and female subgroups. Demographic indicators (such as self-identified gender, education of parents [101], support by meaningful others) were used to control for differences since such measures have been linked to educational outcomes [106]. Note that all the students in this study self-identified as male or female and gender fluid identities were not self-identified. Students’ grades¹ in school subjects (physics, math, chemistry, biology, German, English) were collected to ensure that no differences in academic achievement were present. Also, interest in school subjects was measured with a single item (“How much do you like [subject]?”). This item had a 5-point Likert scale (1: “not at all” to 5: “very much”). Interest in school subjects is an important indicator which accounts for motivational issues and is used here to characterize and compare the sample with respect to gender differences. Performance in the Physics Olympiad was collected as a baseline comparison for prior engagement in the competition. This measure is a number score given by the teacher on the basis of a solution sheet for four homework problems in the Physics Olympiad’s first round (maximum is 40 points). Another important covariate was the support by teachers, parents, and peers with reference to the Physics Olympiad. Especially with respect to gender, consistent differences can be found in the support by teachers and parents [49,107]. Based on prior research where these differences have been found, we developed scales to measure the support by teachers, parents, and peers. Each scale contained 4 items and is measured on a 4-point Likert scale (1: “agree” to 4: “disagree”) with satisfying internal consistencies (see Table I).

In order to teach meaningful content, a content analysis was carried out for the physics curricula of the federal states in Germany as collected from the respective ministries of education. Additionally, the content coverage for the International Physics Olympiad was consulted.² These documents were reviewed with regards to content that

relates to electromagnetic radiation, and more concretely, the content of the intervention. A learning sequence for radiation physics emerged from this. The learning sequence was as follows: waves and oscillations, light as electromagnetic wave, properties of waves, and an experiment on light diffraction. These contents formed the basis for the tasks in the content knowledge (CK) test that was administered pre and post in order to check if there were gender differential effects for learning gains (performance) from the intervention. Though administering tests in informal learning environments is contested [108], we considered that the Physics Olympiad in itself is particularly focused on problem solving performance such that acceptance amongst students will be high and important information would have been lost had we not included a content knowledge test. For item development in this test, a physics book [109] and an online resource (www.leifiphysik.de) were consulted. These tasks were transformed into a multiple-choice assessment with 14 items. Each multiple-choice item had four alternatives, one correct answer and three distractors based on commonly held preconceptions. Most importantly, the items specifically covered the content that was the focus of the intervention. Right after each content item the students were given a confidence scale where they indicated how confident they felt about their answer on a 4-point Likert-scale from “very uncertain” to “very certain.”

VI. SAMPLE

All young women ($N = 31$) that could participate in the next year’s Physics Olympiad (2016) and reached the second round of the 2015’s Physics Olympiad were solicited for participation. Young men were matched with regard to similar performance in the first round of the German Physics Olympiad and were also solicited for participation. However, four of the participants (3 female, 1 male) were invited despite the fact that they did not participate in the German Physics Olympiad prior to the intervention. They were nominated by their respective teachers and were all familiar with the Physics Olympiad. In total, 42% of the invited female students and 50% of the invited male students participated in the intervention. Overall, 30 students took part in this intervention (13 female, 17 males). Data for one student is missing because he was not able to participate in the first survey administration. Thus, complete data are available for 29 students (13 female, 16 males³). The students came from various places in Germany, and the majority did not know each other. All the students who participated in the Physics Olympiad had successfully passed the first round and were working on the

¹Note that in the German school system a grade of 1 is best, whereas a grade of 6 is worst.

²see: <http://ipho.org/syllabus.html>, 12 December 2016.

³Note that the students self-identified as females and males in the Olympiad’s online platform and in the questionnaires administered in the intervention study. Thus, only female and male gender identities are considered in the following.

TABLE II. Sample differences in background variables with respect to students' gender.

Variable	men $M(SD)$	women $M(SD)$	t statistic	p value
Age	16.80 (0.73) a	15.80 (0.96) a	$t(18.23) = 2.78$	<0.05
Support				
Teachers	3.48 (0.48)	3.45 (0.56)	$t(26.88) = 0.17$	0.87
Parents	2.71 (0.82)	2.60 (0.63)	$t(22.33) = 0.41$	0.69
Peers	2.31 (0.62)	2.24 (0.76)	$t(27) = 0.26$	0.79
Grades	1.38 (0.39)	1.55 (0.26)	$t(18.09) = -1.27$	0.22
Performance in Physics Olympiad	28.75 (5.61)	25.65 (9.33)	$t(13.1) = 0.95$	0.36
Interest in school subjects	3.98 (0.58)	3.66 (0.35)	$t(19.34) = 1.75$	0.10
Intent of further participation in Physics Olympiad (pre)	6.50 (1.90)	6.31 (1.93)	$t(25.60) = 0.27$	0.79

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

second rounds' physics problems at the time when the intervention occurred.

Since young women and men differ in the population with regards to background variables such as interest in school subjects or physics performance, we first assessed whether such differences appeared in our sample. This is important because, for example, a different baseline interest might affect the experiences that students make in the intervention and thus limit the conclusions of our analyses. Therefore, the means and standard deviations for young women and men in the background variables are displayed in Table II. Since young men function as a control condition for the effects of young women, both of these groups are contrasted with t tests. No significant gender differences were found in the sample in all but one of the background variables (Table II). The age of the participants was different with respect to gender, an effect that cannot be controlled for in the statistical analysis.⁴ This will be discussed later on.

VII. ANALYSIS

In order to explore possible effects, the data were analyzed with statistical hypothesis tests that detect meaningful effects against their likelihood of occurring by mere chance. We emphasize that our data are not meant to test the science identity theory, but rather to explore possible effects that can be expected to be observed when considering the identity constructs in an intervention in an informal science setting. To address parts of RQ 1 and RQ 2 analysis of variance (ANOVA) was used. The ANOVA compares several means of groups and is robust against violations for the normality assumption. In order to account for the dependent measures over time, a factorial repeated measures ANOVA was used with time and gender as within and between factors, respectively. We used the statistics software R with the package "ez" (version 4.4-0) to perform

⁴Note that the significance disappears when the p -value is adjusted for multiple comparisons, e.g., through Bonferroni correction.

this analysis [110]. Mainly type II sums of squares were used [111]. To report effect sizes a generalized eta squared (η^2) was used that is comparable with the well-known eta squared (η^2) from ANOVA [112]. When we used t tests, an effect size r was calculated. The effect size measure r can be used to characterize the strength of an effect. If r is bigger than 0.30 the effect size is considered of medium size. An effect size of $r > 0.50$ is considered a large effect [110]. Even though the ANOVA is robust against violations of the normality assumption, nonparametric tests were included at times (RQ 1) in order to test the effects without making the normality assumption [110]. Wilcoxon rank sum test compares group means and can be used for repeated measures as well. When the dependent variable was not measured pre and post (RQ 1a), ANCOVA was used to account for other influencing variables. In order to check whether the predictors were independent of the focal categorical predictor variable, i.e., gender, a MANOVA was used. A MANOVA can accommodate for multiple dependent variables and assess the association of a predictor variable with the dependent variables.

Overall, amongst the dependent variables, 9% had missing values. Regression random imputation [113] was used as a means to retain these values for the analyses. In order to impute the values, regression models were fit with gender, age, and competence beliefs in physics as predictors.

VIII. RESULTS

The results section is arranged alongside the RQs: how does physics identity develop for students that participated in the intervention (RQ1), and to what extent does the intervention affect young women's and men's intended and future participation in the Physics Olympiad (RQ2). RQ1 is subdivided into the different identity constructs: interest (RQ 1a), recognition (RQ 1b), and competence (RQ 1c). A similar division is done for RQ 2. Here, the intention for future participation in the Physics Olympiad (RQ 2a), and the actual future participation (RQ 2b) are presented

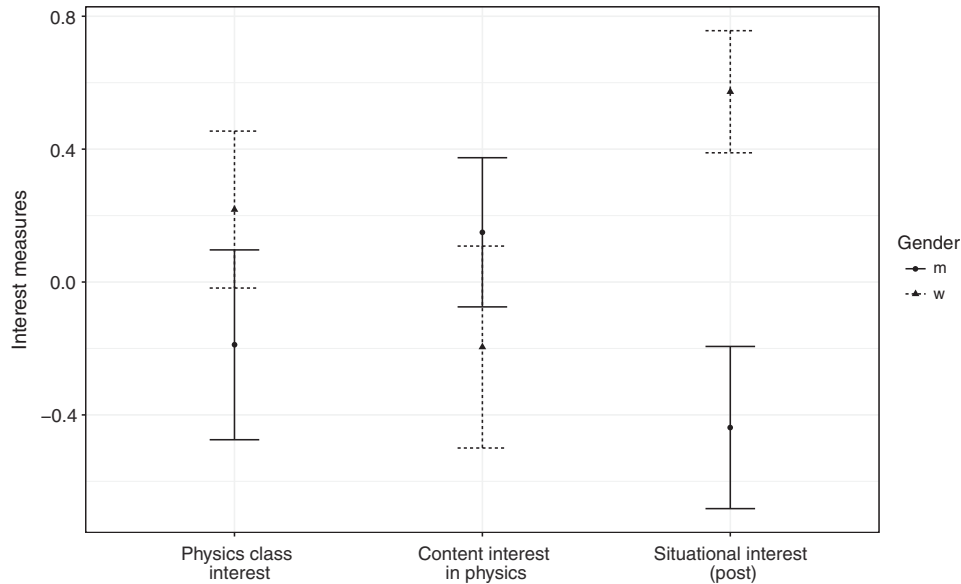


FIG. 3. Means and standard errors for interest measures with regards to gender differences (note that the variables are standardized).

separately. Finally, we also examined gains in content knowledge as a result of the intervention.

A. Interest (RQ 1a)

Figure 3 presents an overview of the *interest* scales. Note the variables are z scored (around grand mean of each variable) in order to highlight differences between the gender groups and account for different response formats of the interest scales. The effects of gender on situational interest (post) is assessed through an ANCOVA. In order to ensure that gender was independent of the other predictors (i.e., content interest in physics and physics class interest) prior to running the ANCOVA, a MANOVA was implemented. The results indicated that the predictors do not depend on gender, $F(2, 27) = 1.69$, $p = 0.202$. In order to better understand the influence of gender on situational interest (post), with physics class interest and content interest in physics controlled for, an ANCOVA was calculated with situational interest (post) as the dependent variable. The model explained 21% ($R_{adj}^2 = 0.21$) of the variance in situational interest (post). Gender had a significant main effect, $\beta = 0.66$, $se = 0.21$, $t = 3.18$, $p < 0.01$, after controlling for the other influences. No other effects were significant.

B. Recognition (RQ 1b)

For recognition there was a main effect for gender in the repeated measures ANOVA, $F(1, 28) = 8.17$, $p < 0.01$, with a medium effect size of $\eta^2 = 0.15$ (see Fig. 4). The young women reported higher recognition in the Physics Olympiad than the young men across both time points. For recognition in physics class the effect for gender was not

significant, $F(1, 28) = 3.54$, $p = 0.070$. The effect for time was also not significant.

C. Competence (RQ 1c)

A repeated-measures ANOVA was used to examine whether competence in the Physics Olympiad was significantly different for women and men before and after the intervention (see Fig. 5). The time main effect was not significant. However, there was a significant interaction effect between gender and time, $F(1, 28) = 5.23$, $p < 0.05$, with a small effect size $\eta^2 = 0.02$. This means that young women improved their competence belief within the Physics Olympiad over time compared to young men. Even though ANOVA is robust with respect to the normality assumption for large samples [114], nonparametrical tests were also used since the variables were not normally distributed and our sample was small. The Wilcoxon signed rank sum test was used to test effects [110]. The change over time for the female students is significant, $V = 45$, $p < 0.01$, with a large effect $r = 1.00$, whereas for the males this effect is nonsignificant, $V = 30$, $p = 0.82$.

For competence in physics class, no main effects were significant. As with competence in the Physics Olympiad, a significant interaction between gender and time was found, $F(1, 28) = 4.87$, $p < 0.05$, with a small effect size of $\eta^2 = 0.07$. Similar to feelings of competence with respect to the Physics Olympiad, the young women, compared to young men, showed a significant increase in their feeling of competence for their physics class after the intervention.

D. Engagement in Physics Olympiad (RQ 2a and 2b)

The participants in the intervention were asked pre and post how likely they would participate in the next year's Physics

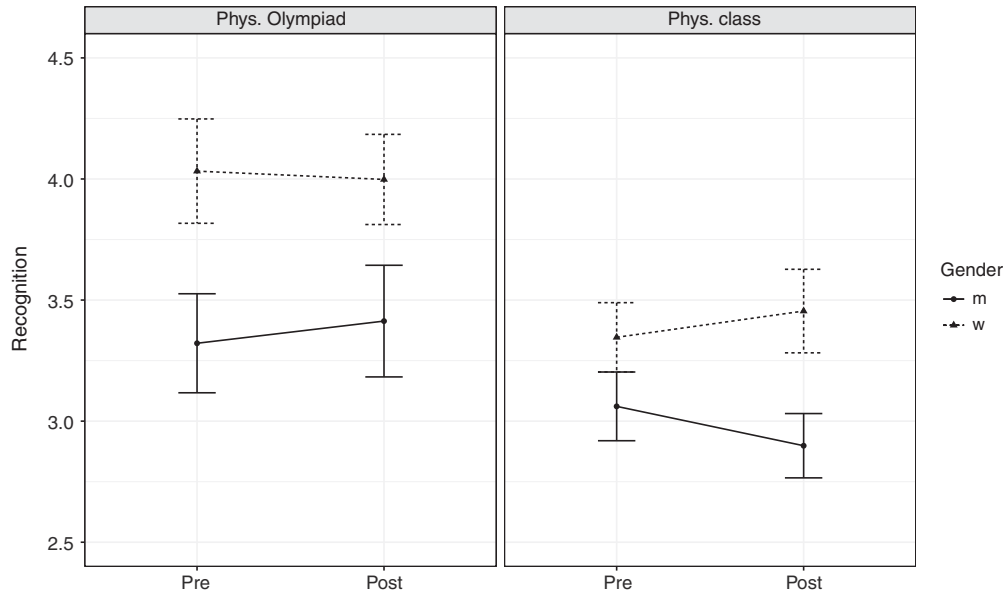


FIG. 4. Means and standard errors for recognition over time with regard to gender.

Olympiad (RQ 2a). In the postmeasurement the male students had a mean of 6.46 ($SD = 1.85$) on the 8-interval scale. The mean thus falls in the interval that indicated a 75% to 87.5% likelihood of participating again in the next year’s Physics Olympiad. The female students had a mean of 6.25 ($SD = 1.66$) which falls into the same interval, 75% to 87.5% likelihood of returning to the next year’s Physics Olympiad. No significant effects for time, gender, or the interaction of both appear in the repeated measures ANOVA.

Considering RQ 2b, in the overall population of competition participants in the Physics Olympiad 2015–2016, 39% of the males and 28% of the females (who were not in their final grade in high school) participated again in the

competition. For the participants in the intervention that also participated in the earlier Physics Olympiad, 62% (10 out of 16) of the young men and 60% (6 out of 10) of the young women participated in the next year’s Physics Olympiad. For the students that did not initially participate in the Physics Olympiad, 33% (1 out of 3) of the young women and none of the young men (0 out of 1) participated in the next year’s Physics Olympiad.

E. Content knowledge

Using the 14 item content knowledge test specific to the intervention topics, we examined whether young women

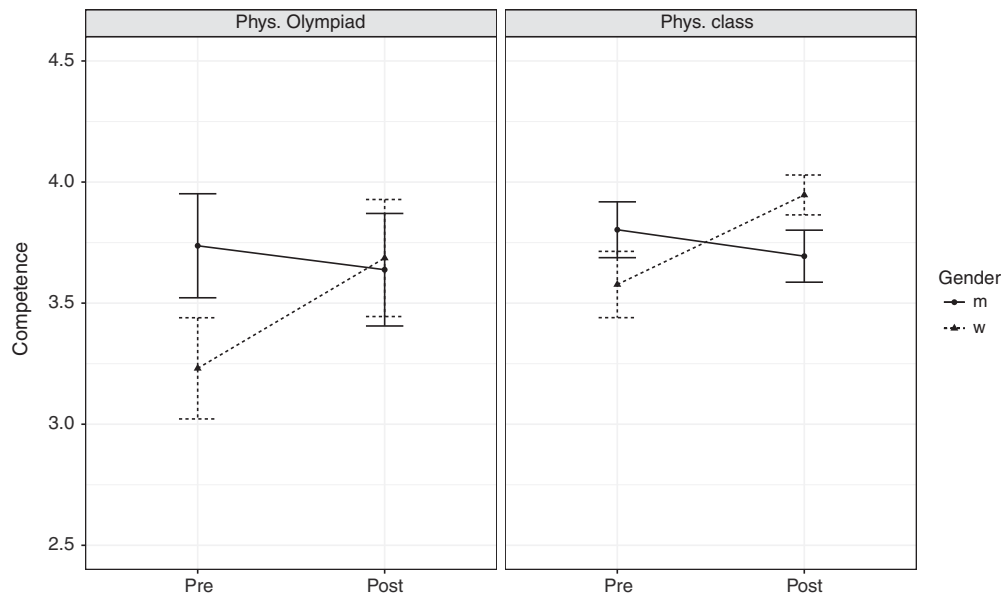


FIG. 5. Means and standard errors for competence over time with regard to gender.

and men had comparable gains in content knowledge and confidence after the intervention. The effects for content knowledge were examined over time and between gender groups with a two-way repeated measures ANOVA. The items in the CK test were solved postintervention both more accurately and with more confidence. The effect for time is significant and large for CK and confidence, $F(1, 23) = 55.41$, $p < 0.001$, $\eta^2 = 0.24$ and $F(1, 23) = 121.18$, $p < 0.001$, $\eta^2 = 0.48$, respectively. No gender differences appeared either in confidence or performance over the intervention.

IX. DISCUSSION

As a means to increase gender equity in informal science learning programs such as the Physics Olympiad, the intervention presented here allowed us to explore the effects that we can expect when designing an intervention in this context grounded in physics identity theory. The intervention aimed to support engagement and physics identity development for young women who participate in the Physics Olympiad in Germany. The results are suggestive of the conclusion that young women were positively affected with regards to physics identity by this intervention while no detrimental effects appeared for young men. For example, young women and men equally gained content knowledge and confidence in their responses in the content knowledge test. This gives us confidence that no gender was at a disadvantage with respect to the content that was taught. Both genders learned equally in terms of conceptual and factual knowledge. With regards to the identity construct interest (RQ 1a) it was shown that the young women were particularly interested in the intervention contents reversing the traditional gender difference observed in the literature of male students being more interested in physics than female students [115] and is consistent with studies that indicate that targeted interventions can enhance young women's interest in physics [67]. Young men were not negatively affected by the intervention with respect to *interest*. Although it cannot be seen in Fig. 3, the young men expressed a high situational interest (post). For example, the unscaled mean of the males is 3.78 (SD = 0.58), which is significantly larger than the mean of the scale (i.e., 3.00), $t(16) = 5.54$, $p < 0.001$. Furthermore, young women responded more positive on the identity construct recognition (RQ 1b) in the Physics Olympiad than young men. We did not expect this pattern to emerge, however, a possible explanation would be that the young women were more affected that they were subject to an intervention in the context of the Physics Olympiad compared to the young men and thus felt more recognized by the Physics Olympiad. Another explanation would be that young women in general need a higher feeling of recognition by the community in order to subscribe to a program such as the Physics Olympiad. This aligns with the finding by Lock, Hazari, and Potvin, namely, that women require a greater recognition

in physics in order to persist [116]. On the other hand, this result resonates with the finding that young women, compared to young men, were sometimes found to score higher in school aspiration and attitudes, such as recognition by others [117]. Recognition does not change for either young women or men. However, the low internal consistency of the scale reduces the validity of any such conclusion. The low internal consistency might stem from the fact that the constructs of "IPhO-team" and "other" are very different for students (see items in Table I). With regards to the third identity construct, competence, the young women improved their competence beliefs towards the Physics Olympiad and the physics classroom throughout the intervention (RQ 1c). The values for competence for young men remained unchanged, $M(\text{pre}) = 3.74$, $SD = 0.89$; $M(\text{post}) = 3.64$, $SD = 0.97$; $t(31.73) = 0.29$, $p = 0.777$, $r = 0.05$.

To back up these findings, we included further measures that relate to positive affiliation with the Physics Olympiad (RQ 2). While there were no gender effects for the reported intent of future participation in the Physics Olympiad (RQ 2a), this can be interpreted as a promising result given the typically high attrition rate of women in the Physics Olympiad. Furthermore, we tracked students' participation in the next year's Physics Olympiad and found an overall higher return rate of the participants in this sample compared to the population (RQ 2b). The return rate for young women in the intervention group was 60% while the return rate for young women in the overall Olympian population was 28%. For the young men, this return rate was also higher for the intervention group (62%) than for the general population of male Olympians (39%).

Given the sample and the size of the intervention the results have to be interpreted cautiously with regard to inferences for the Physics Olympiad population and beyond. First, the high initial motivation of these students points to a potential shortcoming of the intervention design and conclusions, i.e., these students are already more motivated towards physics and the Physics Olympiad than most other students. However, the performance of the participants in the intervention turns out to be not significantly different compared to the performance of the general Olympian population [$M(\text{intervention}) = 27.56$, $SD = 7.25$; $M(\text{Olympiad}) = 24.76$, $SD = 10.4$; $t(33.75) = 1.82$, $p = 0.077$] which indicates that the intervention sample is in some respects similar to the overall Olympian population. Another important aspect to consider is the significant difference in age between females and males. Such effects cannot be ruled out in the repeated measures ANCOVA [118] or any other statistical technique [119] since no information for older females is available. The fact that the young women were younger than young men fuels our hope that such carefully designed interventions yield positive effects for young women, since the age difference likely sets the young women at a disadvantage in terms of initial physics knowledge and experience with

physics. Such a disadvantage did not appear in the intervention. Furthermore, possible effects that the test-taking situation had on students' responses could not be disentangled from the data. Finally, the small sample size results in low statistical power and a potential inflation of type II error rate. Despite this limitation, we still observed significant effects in this study. Our results also align well with prior research and none of the reported effects is contradictory to prior findings. However, it would be beneficial if future research focuses on confirming these results with larger samples.

X. CONCLUSION

We are ultimately interested in adapting the social context of informal science learning programs such as the Physics Olympiad in order to raise gender equity in these programs. The reported results point to some important aspects for further efforts in the Physics Olympiad (and probably other informal science environments) that particularly facilitate physics identity development for young women. The results are in line with findings that indicate that topics which are particularly interesting to female students do not depress the interest of male students [120]. The fact that young women report a significantly higher intervention topic interest is an important finding to further our understanding of how to better address specific issues that concern young women. The learning materials on the case study of Rosalind Franklin and the DNA structural analysis appeared to be particularly motivating for the young women in the sample. However, this one-time intervention seemed to have no effect on the important recognition construct for physics identity as it relates to the Physics Olympiad. The low reliability of the scale suggests adopting better measures for this construct, possibly as in the study in the math context by Cribbs *et al.* [34]. Improving the design to enhance recognition may be fruitful for improving the intervention since recognition has been found to be a key aspect to support young women in physics [27] and was more predictive for math identity

than interest [34]. As such, intervention strategies, such as self-to-prototype matching, can play an important role to improve students' perceived recognition by the physics community [53].

On the basis of the results it can be suggested that science identity theory is a useful framework for designing interventions that increase gender equity in STEM. The pre-post assessment of science identity and the external criterion (i.e., enrollment in next Physics Olympiad) enabled us to explore effects of such an intervention. The results indicate that the participants are more likely to enroll in the next Physics Olympiad compared to the overall Physics Olympiad population, and that some identity constructs (interest, competence, performance) can be affected with such a targeted small-scale intervention. Further qualitative research is necessary in order to understand the effects of the science identity constructs on enrollment and engagement measures to better reveal the mechanisms that lead to identity development in the context of informal science environments. However, we also submit that potential mechanisms for underrepresentation of young women in these programs might remain undetected in the identity framework. Some studies point to differences in underlying motives for participation in such programs, e.g., competitiveness [121], for young women and young men such that the overall framing of these programs might differentially appeal to young women and men. The identity lens enabled us to design an intervention considering multiple facets of physics engagement and explore possibilities for changing the competition environment (in its due constraints set by the international competition).

ACKNOWLEDGMENTS

The authors thank Jeff Nordine for his comments on an earlier version of the manuscript and the reviewers for their thorough comments. Furthermore, in the context of this paper, ESERA supported a research visit of the first author in the PER group at Florida International University through a travel grant.

-
- [1] T. L. Friedman, *The World is Flat* (Farrar, Straus, and Giroux, New York, 2005).
 - [2] A. Schiepe-Tiska *et al.*, in *PISA 2015. Eine Studie zwischen Kontinuität und Innovation*, edited by K. Reiss *et al.* (Waxmann, Münster, New York, 2016), p. 45.
 - [3] X. Chen, *STEM Attrition. College students' paths into and out of STEM fields (NCES 2014-001)* (National Center for Education Statistics, Washington, DC, 2013).
 - [4] E. Smith, Women into science and engineering? Gendered participation in higher education STEM subjects, *Br. Educ. Res. J.* **37**, 993 (2011).
 - [5] J. Handelsman *et al.*, More women in science, *Science* **309**, 1190 (2005).
 - [6] G. Stoet, D. H. Bailey, A. M. Moore, and D. C. Geary, Countries with higher levels of gender equality show larger national sex differences in mathematics anxiety and relatively lower parental mathematics valuation for girls, *PLoS One* **11**, e0153857 (2016).
 - [7] J. R. Campbell, H. Wagner, and H. J. Walberg, in *International Handbook of Giftedness and Talent*, edited by K. Heller, F. Monks, R. Sternberg, and R. Subotnik (Pergamon, Oxford, UK, 2000), p. 523; A. Ziegler and

- H. Stoeger, in *Gifted and Talented Education*, edited by D. Eyre (Routledge, London, 2009), p. 389.
- [8] S. Petersen and P. Wulff, The German Physics Olympiad—identifying and inspiring talents, *Eur. J. Phys.* **38**, 034005 (2017).
- [9] S. M. Reis and J. S. Renzulli, Is there still a need for gifted education? An examination of current research, *Learn. Indiv. Differ.* **20**, 308 (2010).
- [10] A. M. Aljughaiman and A. E. A. Ayoub, The effect of an enrichment program on developing analytical, creative, and practical abilities of elementary gifted students, *J. Educ. Gifted* **35**, 153 (2012); J. Wai, D. Lubinski, C. P. Benbow, and J. H. Steiger, Accomplishment in science, technology, engineering, and mathematics (STEM) and its relation to STEM educational dose: A 25-year longitudinal study, *J. Educ. Psychol.* **102**, 860 (2010); J. L. Wirt, Ph.D. thesis, Seton Hall University, 2011.
- [11] A. X. Feng, J. R. Campbell, and M. A. Verna, The talent development of American Physics Olympians, *Mensa Res. J.* **34**, 47 (2003); F. Oswald, G. Hanisch, and G. Hager, *Wettbewerbe und "Olympiaden". Impulse zur (Selbst-)Identifikation von Begabungen* (Lit, Berlin, 2004).
- [12] J. R. Campbell, Preface, *Int. J. Educ. Res.* **25**, 477 (1996); J. R. Campbell and H. J. Walberg, Olympiad studies: Competitions provide alternatives to developing talents that serve national interests, *Roeper Rev.* **33**, 8 (2011); G. Lind, *Herausforderung, Selbstbestätigung durch Erfolg, Erfahrung der eigenen Grenzen. Eine Befragung ehemaliger Teilnehmer der Physikolympiade* (IPN, Kiel, 2001).
- [13] J. R. Campbell, Gender Inequity Among Academic Olympians Across the Globe: Theoretical Paradigms, *J. Res. Educ.* **12**, 75 (2002).
- [14] A. Steegh, T. N. Höffler, and I. Parchmann (to be published).
- [15] A. Lengfelder and K. A. Heller, German olympiad studies: Findings from a retrospective evaluation and from in-depth interviews. Where have all the gifted females gone?, *J. Res. Educ.* **12**, 86 (2002).
- [16] A. Ziegler and K. A. Heller, in *Munich Studies of Giftedness*, edited by K. A. Heller (Lit, Berlin, 2010), p. 313.
- [17] N. Dasgupta, M. M. Scircle, and M. Hunsinger, Female peers in small work groups enhance women's motivation, verbal participation, and career aspirations in engineering, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4988 (2015).
- [18] J. E. Stets and P. J. Burke, Identity Theory and Social Identity Theory, *Soc. Psycho. Quart.* **63**, 224 (2000).
- [19] J. E. Stake and K. R. Mares, Science enrichment programs for gifted high school girls and boys: Predictors of program impact on science confidence and motivation, *J. Res. Sci. Teach.* **38**, 1065 (2001).
- [20] S. J. Ceci, W. M. Williams, and S. M. Barnett, Women's underrepresentation in science: Sociocultural and biological considerations, *Psychol. Bull.* **135**, 218 (2009).
- [21] A. Bandura, *Self-Efficacy. The Exercise of Control* (W. H. Freeman, New York, 1997).
- [22] L. Archer, J. DeWitt, J. Osborne, J. Dillon, B. Willis, and B. Wong, "Balancing acts": Elementary school girls' negotiations of femininity, achievement, and science, *Sci. Educ.* **96**, 967 (2012); N. W. Brickhouse, P. Lowery, and K. Schultz, What kind of a girl does science? The construction of school science identities, *J. Res. Sci. Teach.* **37**, 441 (2000); M. Varelas, J. Settlage, and F. M. Mensah, Explorations of the structure–agency dialectic as a tool for framing equity in science education, *J. Res. Sci. Teach.* **52**, 439 (2015).
- [23] H. Joas, *The Creativity of Action* (The University of Chicago Press, Chicago, IL, 1996).
- [24] C. Good, A. Rattan, and C. S. Dweck, Why do women opt out? Sense of belonging and women's representation in mathematics, *J. Personality Social Psychol.* **102**, 700 (2012).
- [25] U. Kessels, *Undoing Gender in der Schule. Eine empirische Studie über Koedukation und Geschlechtsidentität im Physikunterricht* (Juventa, Weinheim, München, 2002).
- [26] D. Sekaquaptewa and M. Thompson, Solo status, stereotype threat, and performance expectancies: Their effects on women's performance, *J. Exp. Soc. Psychol.* **39**, 68 (2003).
- [27] H. B. Carlone and A. Johnson, Understanding the science experiences of successful women of color: Science identity as an analytic lens, *J. Res. Sci. Teach.* **44**, 1187 (2007).
- [28] Z. Hazari, G. Sonnert, P. Sadler, and M.-C. Shanahan, Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study, *J. Res. Sci. Teach.* **47**, 978 (2010).
- [29] A. Sfard and A. Prusak, Telling identities: In search of an analytic tool for investigating learning as a culturally shaped activity, *Educ. Res.* **34**, 14 (2005).
- [30] W. Meeus, J. Iedema, M. Helsen, and W. Vollebergh, Patterns of adolescent identity development: Review of literature and longitudinal analysis, *Dev. Rev.* **19**, 419 (1999).
- [31] *Advances in Identity Theory and Research*, edited by P. J. Burke, T. J. Owens, R. T. Serpe, and P. A. Thoits (Kluwer Academic, New York, 2003).
- [32] J. P. Gee, *An Introduction to Discourse Analysis: Theory and Method* (Routledge, New York, 1999).
- [33] J. M. Kane, Young African American children constructing academic and disciplinary identities in an urban science classroom, *Sci. Educ.* **96**, 457 (2012).
- [34] J. D. Cribbs, Z. Hazari, G. Sonnert, and P. M. Sadler, Establishing an explanatory model for mathematics identity, *Child Dev.* **86**, 1048 (2015).
- [35] A. Maltese and R. H. Tai, Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students, *Sci. Educ.* **95**, 877 (2011); R. H. Tai, C. Qi Liu, A. Maltese, and X. Fan, Career choice. Planning early for careers in science, *Science* **312**, 1143 (2006).
- [36] S. Hidi and K. A. Renninger, The four-phase model of interest development, *Educ. Psychol.* **41**, 111 (2006).
- [37] O. Köller, J. Baumert, and K. Schnabel, Does interest matter? The relationship between academic interest and achievement in mathematics, *J. Res. Math. Educ.* **32**, 448 (2001).
- [38] M. M. Bleeker and J. E. Jacobs, Achievement in math and science: Do mothers' beliefs matter 12 years later?, *J. Educ. Psychol.* **96**, 97 (2004); S. L. Turner, J. C.

- Steward, and R. T. Lapan, Family factors associated with sixth-grade adolescents' math and science career interests, *Career Develop. Quarterly* **53**, 41 (2004).
- [39] K. Bussey and A. Bandura, Social cognitive theory of gender development and differentiation, *Psychol. Rev.* **106**, 676 (1999).
- [40] F. Pajares and L. Graham, Self-efficacy, motivation constructs, and mathematics performance of entering middle school students, *Contemp. Educ. Psychol.* **24**, 124 (1999).
- [41] H. W. Marsh, D. Chessor, R. Craven, and L. Roche, The effects of gifted and talented programs on academic self-concept: The big fish strikes again, *Am. Educ. Res. J.* **32**, 285 (1995).
- [42] H. W. Marsh, K.-T. Hau, and C.-K. Kong, Multilevel causal ordering of academic self-concept and achievement: Influence of language of instruction (English compared with Chinese) for Hong Kong students, *Am. Educ. Res. J.* **39**, 727 (2002).
- [43] C. Cass *et al.*, *Examining the Impact of Mathematics Identity on the Choice of Engineering Careers for Male and Female Students* (IEEE, Rapid City, SD, 2011).
- [44] Z. Hazari, C. Cass, and C. Beattie, Obscuring power structures in the physics classroom: Linking teacher positioning, student engagement, and physics identity development, *J. Res. Sci. Teach.* **52**, 735 (2015).
- [45] J. S. Eccles, Gendered educational and occupational choices: Applying the Eccles *et al.* model of achievement-related choices., *Int. J. Behav. Dev.* **35**, 195 (2011).
- [46] D. Baker and R. Leary, Letting girls speak out about science, *J. Res. Sci. Teach.* **40**, 176 (2003).
- [47] J. S. Eccles, Understanding women's educational and occupational choices, *Psychol. Women Q.* **18**, 585 (1994).
- [48] G. Jones, A. Howe, and M. Rua, Gender differences in students' experiences, interests, and attitudes toward science and scientists, *Sci. Educ.* **84**, 180 (2000).
- [49] L. Hoffmann, P. Häußler, and M. Lehrke, *Die IPN-Interessenstudie Physik* (IPN, Kiel, 1998), Vol. 158.
- [50] D. Lubinski and C. P. Benbow, Study of mathematically precocious youth after 35 years: Uncovering antecedents for the development of math-science expertise, *Perspect. Psychol. Sci.* **1**, 316 (2006); D. J. Prediger, Dimensions underlying Holland's hexagon: Missing link between interests and occupations?, *J. Vocat. Behav.* **21**, 259 (1982).
- [51] N. W. Brickhouse and J. T. Potter, Young women's scientific identity formation in an urban context, *J. Res. Sci. Teach.* **38**, 965 (2001).
- [52] P. Labudde, W. Herzog, M. P. Neuenschwander, E. Violi, and C. Gerber, Girls and physics: teaching and learning strategies tested by classroom interventions in grade 11, *Int. J. Sci. Educ.* **22**, 143 (2000); H. Stadler, R. Duit, and G. Benke, Do boys and girls understand physics differently?, *Phys. Educ.* **35**, 417 (2000).
- [53] B. Hannover and U. Kessels, Self-to-prototype matching as a strategy for making academic choices. Why high school students do not like math and science, *Learn. Instr.* **14**, 51 (2004).
- [54] C. Good, J. Aronson, and J. A. Harder, Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses, *J. Appl. Dev. Psychol.* **29**, 17 (2008).
- [55] R. Taconis and U. Kessels, How choosing science depends on students' individual fit to 'Science Culture', *Int. J. Sci. Educ.* **31**, 1115 (2009).
- [56] U. Kessels and B. Hannover, in *Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung*, edited by J. Doll and M. Prenzel (Waxmann, Münster, 2004), p. 398; G. C. Marchand and G. Taasoobshirazi, Stereotype threat and women's performance in physics, *Int. J. Sci. Educ.* **35**, 3050 (2013).
- [57] S.-J. Leslie, A. Cimpian, M. Meyer, and E. Freeland, Expectations of brilliance underlie gender distributions across academic disciplines, *Science* **347**, 262 (2015).
- [58] A. B. Diekmann, E. R. Brown, A. M. Johnston, and E. K. Clark, Seeking congruity between goals and roles: a new look at why women opt out of science, technology, engineering, and mathematics careers, *Psychol. Sci.* **21**, 1051 (2010).
- [59] A. Madsen, S. B. McKagan, and E. C. Sayre, Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap?, *Phys. Rev. ST Phys. Educ. Res.* **9**, 020121 (2013).
- [60] D. F. Halpern, C. P. Benbow, D. C. Geary, R. C. Gur, J. S. Hyde, and M. A. Gernsbacher, The science of sex differences in science and mathematics, *Psychol. Sci. Publ. Interest* **8**, 1 (2007).
- [61] S. Chambers and T. Andre, Gender, prior knowledge, interest, and experience in electricity and conceptual change text manipulations in learning about direct current, *J. Res. Sci. Teach.* **34**, 107 (1997); Z. Hazari, P. Sadler, and R. H. Tai, Gender differences in the high school and affective experiences of introductory college physics students, *Phys. Teach.* **46**, 423 (2008).
- [62] U. Kessels, M. Rau, and B. Hannover, What goes well with physics? Measuring and altering the image of science, *Br. J. Educ. Psychol.* **76**, 761 (2006).
- [63] R. Driver, P. Newton, and J. Osborne, Establishing the norms of scientific argumentation in classrooms, *Sci. Educ.* **84**, 287 (2000).
- [64] T. Schmader, Gender identification moderates stereotype threat effects on women's math performance, *J. Exp. Soc. Psychol.* **38**, 194 (2002).
- [65] J. A. Fredricks and J. S. Eccles, Children's competence and value beliefs from childhood through adolescence: Growth trajectories in two male-sex-typed domains, *Develop. Psychol.* **38**, 519 (2002); K. Kling, J. S. Hyde, C. Showers, and B. Buswell, Gender differences in self-esteem: A meta-analysis, *Psychol. Bull.* **125**, 470 (1999).
- [66] P. Häußler and L. Hoffmann, An intervention study to enhance girls' interest, self-concept, and achievement in physics classes, *J. Res. Sci. Teach.* **39**, 870 (2002).
- [67] R. Berger, *Z. Didaktik der Naturwissenschaften* **8**, 119 (2002).
- [68] J. Jovanovic and S. Steinbach King, Boys and Girls in the Performance-Based Science Classroom: Who's Doing the Performing?, *Am. Educ. Res. J.* **35**, 477 (1998);

- D. H. Palmer, Student interest generated during an inquiry skills lesson, *J. Res. Sci. Teach.* **46**, 147 (2009).
- [69] L. Aguilar, G. Walton, and C. Wieman, Psychological insights for improved physics teaching, *Phys. Today* **67**, No. 5, 43 (2014).
- [70] G. L. Cohen, J. Garcia, V. Purdie-Vaughns, N. Apfel, and P. Brzustoski Recursive processes in self-affirmation: Intervening to close the minority achievement gap, *Science* **324**, 400 (2009); G. M. Walton and G. L. Cohen, A brief social-belonging intervention improves academic and health outcomes of minority students, *Science* **331**, 1447 (2011).
- [71] T. Hodapp and Z. Hazari, Women in Physics: Why so few?, *APS News* **24**, 8 (2015).
- [72] L. Breger, *From instinct to identity, The Development of Personality* (Prentice Hall, Englewood Cliffs, NJ, 1974); B. Hannover, Die Bedeutung des pubertären Reifestatus für die Herausbildung informeller Interaktionsgruppen in koedukativen Klassen und in Mädchenschulklassen, *Z. Pädagogische Psychologie* **11**, 3 (1997).
- [73] E. H. Erikson, *Childhood and Society* (Norton, New York, 1963).
- [74] J. E. Marcia, in *Handbook of Adolescent Psychology*, edited by J. Adelson (Wiley, New York, 1980), p. 159.
- [75] Z. Hazari, E. Brewe, R. M. Goertzen, and T. Hodapp, *Phys. Teach.* **55**, 96 (2017); R. Ivie, R. Cuzjko, and K. Stowe, Women physicists speak. The 2001 International Study of Women in Physics, 2001, <http://www.aip.org/statistics/trends/gendertrends.html>; R. Ivie and S. Guo, *Woman Physicists Speak Again, 2006* (AIP Pub., New York, 2006).
- [76] C. M. Steele, A threat in the air: How stereotypes shape intellectual identity and performance, *Am. Psychol.* **52**, 613 (1997).
- [77] G. Stoet and D. C. Geary, The gender-equality paradox in science, technology, engineering, and mathematics education, *Psychol. Sci.* **29**, 581 (2018).
- [78] A. Krapp, Entwicklung und Förderung von Interessen im Unterricht, *Psychologie in Erziehung und Unterricht* **44**, 185 (1998).
- [79] J. S. Eccles, B. L. Barber, and D. M. Jozefowicz, Linking gender to educational, occupational, and recreational choices. Applying the Eccles, *et al.* model of achievement-related choices, 1997, Address given at Janet Spence's Festschrift; M. G. Jones and J. Wheatley, *J. Res. Sci. Teach.* **27**, 861 (1990); E. Seymour and N. Hewitt, *Talking about Leaving. Factors Contributing to High Attrition Rates among Science, Mathematics, and Engineering Undergraduate Majors* (Boulder, CO, 1994).
- [80] P. H. Miller, S. V. Rosser, J. Benigno, and M. L. Zieseniss, A desire to help others: Goals of high-achieving female science undergraduates, *Women's Studies Quarterly* **28**, 128 (2000); P. H. Miller, J. Slawinski Blessing, and S. Schwartz, Gender differences in high-school students' views about science, *Int. J. Sci. Educ.* **28**, 363 (2006).
- [81] S. Sjøberg and C. Schreiner, *The ROSE Project. Overview and Key Findings* (University of Oslo, Oslo, 2010).
- [82] N. Holstermann and S. Bögeholz, Interesse von Jungen und Mädchen an naturwissenschaftlichen Themen am Ende der Sekundarstufe I, *Z. Didaktik der Naturwissenschaften* **13**, 71 (2007).
- [83] H. Stadler, R. Duit, and G. Benke, Do boys and girls understand physics differently?, *Phys. Educ.* **35**, 417 (2000).
- [84] R. Wodzinski, in *Physikdidaktik. Theorie und Praxis*, edited by E. Kircher, R. Girwidz, and P. Häußler (Springer, Heidelberg, Berlin, New York, 2007), p. 559.
- [85] G. Braun, D. Tierney, and H. Schmitzer, How Rosalind Franklin discovered the helical structure of DNA: Experiments in diffraction, *Phys. Teach.* **49**, 140 (2011).
- [86] M. Inzlicht and T. Ben-Zeev, A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males, *Psychol. Sci.* **11**, 365 (2000).
- [87] L. Hoffmann, P. Häußler, and S. Peters-Haft, *An den Interessen von Mädchen und Jungen orientierter Physikunterricht* (IPN, Kiel, 1997).
- [88] P. Potvin and A. Hasni, Interest, motivation and attitude towards science and technology at K-12 levels: a systematic review of 12 years of educational research, *Stud. Sci. Educ.* **50**, 85 (2014).
- [89] L. H. Parker and L. J. Rennie, Teachers' implementation of gender-inclusive instructional strategies in single-sex and mixed-sex science classrooms, *Int. J. Sci. Educ.* **24**, 881 (2002).
- [90] M. Raffaelli and E. Duckett, "We were just talking ...": Conversations in early adolescence, *J. Youth Adolescence* **18**, 567 (1989).
- [91] S. Baron-Cohen, Phi Kappa Phi Forum (2005).
- [92] R. Josselson, *Revising Herself. The Story of Women's Identity from College to Midlife* (Oxford University Press, New York, 1996).
- [93] B. Davies and R. Harré, Positioning: The discursive production of selves, *J. Theory Soc. Behav.* **20**, 43 (1990).
- [94] B. J. Drury, J. O. Siy, and S. Cheryan, When do female role models benefit women? The importance of differentiating recruitment from retention in STEM, *Psychological Inquiry* **22**, 265 (2011).
- [95] D. M. Marx and J. S. Roman, Female role models: Protecting women's math test performance, *Personality Soc. Psychol. Bull.* **28**, 1183 (2002).
- [96] D. E. Meltzer and R. K. Thornton, Resource letter ALIP-1: Active-learning instruction in physics, *Am. J. Phys.* **80**, 478 (2012).
- [97] F. Reif, Millikan Lecture 1994: Understanding and teaching important scientific thought processes, *Am. J. Phys.* **63**, 17 (1995).
- [98] J. F. Bell, Investigating gender differences in the science performance of 16-year-old pupils in the UK, *Int. J. Sci. Educ.* **23**, 469 (2001); Z. Hazari, R. H. Tai, and P. Sadler, Gender differences in introductory university physics performance: The influence of high school physics preparation and affective factors, *Sci. Educ.* **91**, 847 (2007).
- [99] D. T. Burkam, V. E. Lee, and B. A. Smerdon, Gender and science learning early in high school: Subject matter and laboratory experiences, *Am. Educ. Res. J.* **34**, 297 (1997); M. P. Freedman, The influence of laboratory instruction on science achievement and attitude toward science

- across gender differences, *J. Women Minorities Sci. Engin.* **8**, 191 (2001).
- [100] A. Krapp, in *Motivational Psychology of Human Development*, edited by J. Heckhausen (Elsevier, Amsterdam, 2000), p. 109.
- [101] A. Frey *et al.*, *PISA 2006 Skalenhandbuch. Dokumentation der Erhebungsinstrumente* (Waxmann, Münster, New York, München, Berlin, 2009).
- [102] S. Fechner, *Effects of Context-Oriented Learning on Student Interest and Achievement in Chemistry Education* (Logos, Berlin, 2009), Vol. 95.
- [103] R. Eisinga, M. t. Grotenhuis, and B. Pelzer, The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown?, *Int. J. Public Health* **58**, 637 (2013).
- [104] T. Rabe, C. Meinhardt, and O. Krey, Entwicklung eines Instruments zur Erhebung von Selbstwirksamkeitserwartungen in physikdidaktischen Handlungsfeldern, *Z. Didaktik der Naturwissenschaften* **18** (2012).
- [105] L. Bergkvist, Appropriate use of single-item measures is here to stay, *Mark. Lett.* **26**, 245 (2015).
- [106] PISA, ed., *PISA 2006 in Deutschland, Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (Waxmann, Münster, New York, München, Berlin, 2008).
- [107] D. Urhahne, L. H. Ho, I. Parchmann, and S. Nick, Attempting to predict success in the qualifying round of the International Chemistry Olympiad, *High Ability Studies* **23**, 167 (2012).
- [108] P. Bell, B. Lewenstein, A. W. Shouse, and M. A. Feder, *Learning Science in Informal Environments. People, Places, and Pursuits* (National Academies Press, Washington, DC, 2009).
- [109] P. Tipler, *Physics For Scientists and Engineers* (W. H. Freeman, New York, 2004).
- [110] A. Field and J. Miles, *Discovering Statistics Using R* (SAGE Publications, Thousand Oaks, CA, 2012).
- [111] D. Navarro, Learning Statistics with R. A tutorial for psychology students and other beginners, version 0.5, 2015, <http://www.fon.hum.uva.nl/paul/lot2015/Navarro2014.pdf>.
- [112] R. Bakeman, Recommended effect size statistics for repeated measures designs, *Behav. Res. Meth. Instrum. Comput.* **37**, 379 (2005).
- [113] A. Gelman and J. Hill, *Data Analysis Using Regression and Multilevel/Hierarchical Models. Analytical Methods for Social Research* (Cambridge University Press, New York, 2007).
- [114] M. Eid, M. Gollwitzer, and M. Schmitt, *Statistik und Forschungsmethoden. Lehrbuch* (Beltz, Weinheim, Basel, 2010).
- [115] J. Parkinson, D. Hendley, H. Tanner, and A. Stables, Pupils' attitudes to science in key stage 3 of the national curriculum: A study of pupils in south wales, *Res. Sci. Technol. Educ.* **16**, 165 (1998); J. Watson, A. Dawson, and S. McEwen, Sixth form a level students' perceptions of the difficulty, intellectual freedom, social benefit and interest of science and arts subjects, *Res. Sci. Technol. Educ.* **12**, 43 (1994).
- [116] R. M. Lock, Z. Hazari, and G. Potvin, *Physics Career Intentions. The Effect of Physics Identity, Math Identity, and Gender* (PER Conference series, Philadelphia, PA, 2012).
- [117] T. Rampino and M. Taylor, *Gender Differences in Educational Aspirations and Attitudes* (Institute for Social and Economic Research, Colchester, UK, 2013).
- [118] J. Cohen and P. Cohen, *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 2nd ed. (Erlbaum, Hillsdale, NJ, 1983).
- [119] G. A. Miller and J. P. Chapman, Misunderstanding analysis of covariance, *Journal of abnormal psychology* **110**, 40 (2001).
- [120] P. Häußler and L. Hoffmann, Physikunterricht - an den Interessen von Mädchen und Jungen orientiert, *Unterrichtswissenschaft* **23**, 107 (1995).
- [121] T. Abernathy and R. Vineyard, Academic competitions in science, *The Clearing House* **74**, 269 (2001); R. L. Trivers, *Sexual Selection and the Descent of Man: 1871–1971* (Aldine, Chicago, 1972), p. 136.