

RESEARCH

Open Access



Linking out-of-school online mentoring and in-school STEM clubs to support girls in STEM: a pretest–posttest control group study

Heidrun Stoeger^{1*}, Claudia Uebler¹, Sonja Bayer¹, Kathrin J. Emmerdinger¹ and Albert Ziegler²

Abstract

Background Despite extensive efforts, females remain underrepresented in STEM fields. Out-of-school programs have demonstrated the potential to address this problem. However, to ensure sustainability and maximize the impact of these programs, researchers emphasize the importance of combining out-of-school offerings with school offerings and investigating the effectiveness of such combined programs and their success criteria (Dahn et al., 2023). We developed a combined program that integrates an online mentoring program for girls, which adheres to key principles of effective mentoring and has been proven effective in previous studies, with offerings at school (i.e., teacher-led STEM clubs). In planning this combined program, we incorporated essential bridging principles for collaboration between out-of-school and school settings (Fallik et al., 2013). We investigated (a) whether the positive effects of the exclusively out-of-school online mentoring program identified in previous studies could be replicated, (b) whether participation in the combined program yielded more positive effects than participation in the mentoring-only program, and (c) which variables are related to the effectiveness of the combined program.

Results We employed a pretest–posttest control group design, comparing four groups: girls who exclusively participated in the out-of-school online mentoring program ($n=212$) and girls in the combined program ($n=323$) as well as a control group of girls ($n=1234$) and boys ($n=1501$) who did not participate in any of the programs. Participants were weighted using propensity score estimation based on age, school type, and baseline values of the evaluation variables (measured at pretest). Both programs positively impacted participants' elective intentions in STEM, engagement in STEM activities, confidence in their STEM abilities, and STEM grades. The combined program did not generally demonstrate greater effectiveness than the mentoring-only program. However, the effectiveness of the combined program was related to the degree of bridging between the out-of-school and in-school offerings, the quality of the combined program, and the perceived value of STEM among the participants.

Conclusion This study is the first to demonstrate that combining out-of-school online mentoring by female STEM experts and teacher-led STEM clubs can effectively promote girls in STEM. However, several success criteria must be considered to realize added benefits compared to exclusively out-of-school online mentoring.

*Correspondence:
Heidrun Stoeger
heidrun.stoeger@ur.de

Full list of author information is available at the end of the article

Keywords Online mentoring, Bridging in-school and out-of-school STEM education, Female participation in STEM, Confidence in STEM abilities, STEM grades, Elective intentions in STEM, STEM activities, Teacher-led STEM clubs, Control group, Pretest–posttest

Background

Women remain underrepresented in many degree programs and professions in science, technology, engineering, and mathematics (STEM) fields (German Federal Statistical Office, 2024; National Center for Education Statistics, 2024). To address this issue, more and more support programs are being developed that ideally start in high school at the latest, before interest in STEM wanes during adolescence (for reviews, see Prieto-Rodriguez et al., 2020; Sáinz et al., 2022; van den Hurk et al., 2019). Out-of-school programs that emphasize interest-driven learning, foster supportive relationships with peers with similar backgrounds and interests, and provide information about future career possibilities have shown great promise (Dahn et al., 2023; Guenaga et al., 2022; Lawlor et al., 2020; Prieto-Rodriguez et al., 2020; Sáinz et al., 2022; Sasson, 2019; Stoeger et al., 2013, 2016, 2023). However, even though these programs yield positive outcomes, they lose momentum if they do not establish connections with other resources or settings, such as additional out-of-school programs, home environments, communities, and schools (After School Alliance, 2021).

To ensure sustainability and enhance the impact of out-of-school offerings—especially with a focus on equity—researchers advocate establishing connections between out-of-school experiences and other educational settings, such as schools (Dahn et al., 2023; Ito et al., 2013, 2020; Sonnert et al., 2025). Therefore, our study aimed to investigate whether systematic connections between a successfully evaluated out-of-school program supporting girls in STEM and school offerings can lead to a more significant impact. As an out-of-school offering, we selected an online mentoring program for female students in grades five to 13, in which women from STEM fields serve as mentors. Additionally, the program facilitates networking among numerous girls and mentors interested in STEM through a secure online platform, allowing them to explore STEM careers, collaborate on interest-driven STEM projects, and engage in meaningful discussions. We combined this program with in-school teacher-led STEM clubs for mentees and planned the cooperation of out-of-school and in-school offerings according to basic bridging principles known from the literature (Fallik et al., 2013).

Out-of-school STEM promotion for girls through (online) mentoring

Prototypically, mentoring is a relatively stable dyadic relationship between an experienced mentor and a less

experienced mentee. The relationship is characterized by mutual trust and goodwill. Its goal is to promote the mentee's learning, development, and advancement (Ziegler, 2009). Research has shown that mentoring can be an effective strategy to support girls and young women in STEM, especially when the mentors are women working or majoring in STEM fields (Canaan & Mouganie, 2023; Dennehy & Dasgupta, 2017; Guenaga et al., 2022; Hernandez et al., 2017; Stoeger et al., 2013, 2016, 2023). For example, Dennehy and Dasgupta (2017) randomly assigned 150 incoming female engineering students to a female mentor, a male mentor, or no mentor for one year. Senior students who were trained as mentors met monthly with their mentees, who were surveyed before mentor assignment, at midyear, and year-end. Only female mentors, but not male mentors, protected mentees against declines in their confidence in their engineering abilities and their sense of social belonging in engineering. Female mentors also protected mentees from increased feelings of threat and thoughts of switching to another major. Moreover, only mentees with female mentors maintained stable intentions to pursue advanced engineering degrees after college. In contrast, the intentions of students with male mentors or no mentors declined over the first year.

While most research has focused on how STEM mentoring can support young women in their early college years (e.g., Dennehy & Dasgupta, 2017; Hernandez et al., 2017), comparatively little research has been conducted with younger age groups in middle or high school. However, support programs for this age group are critical for two reasons. First, interest in STEM tends to decline during adolescence (for reviews, see Prieto-Rodriguez et al., 2020; Sáinz et al., 2022; van den Hurk et al., 2019). Second, most girls have not yet made definitive career decisions. Thus, unlike in higher education settings, there is still an opportunity to engage with those who might otherwise not consider STEM degree programs and careers.

Two exceptions that focus on younger age groups are the Spanish group mentoring program *Inspira STEAM* (Guenaga et al., 2022) and the Germany-wide online mentoring program *CyberMentor* (Stoeger et al., 2013, 2016, 2023). The *Inspira STEAM* program targets students aged 10–12 years. During six group mentoring sessions led by mentors with professional backgrounds in STEM, small groups of students learn about the significance of STEM in their environment, explore STEM professions, and discuss their connections to societal challenges. They also examine gender stereotypes

and the contributions of historical, current, and local women in STEM. Guenaga et al. (2022) surveyed program participants ($n=190$) and a waitlist control group of students attending the same schools ($n=107$) before and after program participation. Program participants showed increased interest and confidence in technology. Although girls still reported lower values than boys, they showed greater increases from pre- to post-measures, indicating a more substantial program impact for girls than for boys.

In the Germany-wide online mentoring program CyberMentor, girls aged 10–18 receive a personal mentor, either a female STEM professional or a woman majoring in a STEM subject. Participants commit to at least one year of participation, with weekly mentor–mentee exchanges of about 30 min. Communication takes place on a secure online mentoring platform, providing networking opportunities with up to 800 female mentors and 800 other mentees. In an initial evaluation study by Stoeger and colleagues (Stoeger et al., 2013), 208 out of 312 girls who had registered for the CyberMentor program were randomly assigned to the intervention group, while the remaining 104 girls formed a waitlist control group of comparably interested girls, who were admitted to the program one year later. Both groups completed surveys before random assignment to the treatment or control group, after six months into the program, and at the end of one program year. Compared to the control group, participating girls showed increased STEM activities, increased knowledge about STEM majors and careers, and stronger intentions to pursue a STEM track or course at school, a university STEM major, or a STEM career. Furthermore, program participants' confidence in their STEM abilities and STEM competencies remained stable, whereas the control group showed a decline. In a follow-up study (Stoeger et al., 2016), program participants ($n=789$) were compared with a waitlist control group of comparably interested girls ($n=448$) and randomly selected samples of same-aged girls ($n=663$) and boys ($n=841$). Across the three measurement points, participating girls showed greater increases in certainty about career goals than all three comparison groups and more substantial increases in STEM activities than the waitlist control group and the female control group. Further studies identified success criteria of the CyberMentor program, such as mentees' degree of STEM communication and networking behavior on the mentoring platform (Stoeger et al., 2021b, 2016, 2017b; Uebler et al., 2023) as well as the quality of their mentoring relationship (Stoeger et al., 2019). In addition, a long-term follow-up study (Stoeger et al., 2023) compared the university degree and career choices of former mentees ($n=410$) with their age cohort in Germany ($n=252,737$). Compared to their age cohort, former CyberMentor

mentees were more than twice as likely to have chosen a STEM career (23.9% vs. 51.2%).

In addition to specific mentoring programs, mentoring is frequently incorporated into short-term out-of-school workshops or longer-term research projects aimed at inspiring high school girls to pursue STEM (Dick et al., 2024; Jethwani et al., 2017; Lawlor et al., 2020; Martín-Peciña et al., 2025; Sasson, 2019). Overall, research highlights the benefits of mentoring in supporting girls and young women in STEM. However, to bring about positive changes, mentoring programs need to adhere to specific quality standards and success criteria (DuBois et al., 2002, 2011; Eby et al., 2008; Luo & Stoeger, 2023; Raposa et al., 2019; Stoeger et al., 2021a). For example, female mentors currently studying STEM subjects or working in STEM professions are particularly effective at encouraging girls to become interested in STEM fields (Dasgupta, 2011; Dennehy & Dasgupta, 2017; Stout et al., 2011). As positive role models, these mentors can help dismantle stereotypes, enhance self-concept, foster identification with STEM, and increase interest in and the likelihood of pursuing STEM careers (Dasgupta, 2011; Gladstone & Cimpian, 2021; González-Pérez et al., 2020; Lawner et al., 2019; Stout et al., 2011). Ideally, female mentors serve as role models and take on additional important roles. They can facilitate interest-driven learning, answer STEM-related questions, provide insights about courses and careers, assist with socio-emotional challenges, and act as brokers for learning opportunities, social connections, or valuable information sources (Kupersmidt et al., 2018; Santo et al., 2019).

Furthermore, to be successful, mentoring programs should meet specific structural requirements. These include ensuring sufficient duration and frequency of contact between mentors and mentees, as well as providing a conducive environment for effective mentoring (DuBois et al., 2002, 2011; Eby et al., 2013; Garringer et al., 2015; Grossman et al., 2012; Grossman & Rhodes, 2002). Research also indicates that matching mentees with suitable mentors and providing training and support for both mentees and mentors from program staff are key quality factors that contribute to successful mentoring outcomes (Blake-Beard et al., 2011; Keller & DuBois, 2021; Lyons & Edwards, 2022; McQuillin & Lyons, 2021; Stelter et al., 2021). Successful mentoring programs are also characterized by enabling supportive relationships between mentor and mentee, as well as with other individuals with similar backgrounds and interests. Networking with peers and role models of similar status has proven to be notably beneficial (Dasgupta, 2011; Hopp et al., 2020; Stoeger et al., 2021b, 2017b).

Online mentoring is particularly effective at integrating essential quality features (Emmerdinger et al., 2024; Stoeger et al., 2021b, 2023). The geographical and

scheduling flexibility offered by online mentoring allows for the recruitment of women studying or working in STEM fields to participate as mentors, despite the overall low participation of women in these fields. Additionally, secure online platforms create an ideal environment for mentoring, facilitating synchronous and asynchronous communication, such as chats, emails, video calls, and forums, thus allowing mentors, despite their demanding schedules and challenging careers, to engage with their mentees regularly and consistently. Moreover, online platforms facilitate networking opportunities with other female mentees and mentors interested in STEM, collaboration on STEM projects, and flexible, needs-based training and assistance for mentees and mentors.

Connecting out-of-school (online) mentoring with school offerings to increase the effectiveness of STEM promotion for girls

Most existing out-of-school (online) mentoring programs supporting girls in STEM are exclusively out-of-school offerings. In the few exceptions where collaboration with schools occurs, it is usually only at a logistical level (Ziegler et al., 2025a, b). For example, in the STEAM Inspira program (Guenaga et al., 2022), the six group mentoring sessions are conducted on-site at the participating schools during school hours. A coordinating team at each school organizes the time and location of the sessions; however, the sessions themselves are led exclusively by out-of-school STEM mentors. In another one-year after-school program (Finkel, 2017) in which mentoring elements are incorporated into STEM enrichment programs offered through university–school partnerships, undergraduate student mentors support high school students in identifying, planning, and carrying out community-based STEM research projects. Undergraduate mentors and participating students meet weekly at the schools, and teacher liaisons assist with on-site logistical challenges such as finding meeting spaces, ensuring laboratory safety, and securing the necessary materials. Potentially, such collaborations have a more positive impact than exclusively out-of-school mentoring and lose momentum less quickly (After School Alliance, 2021; Fallik et al., 2013). However, only descriptive data from post-program surveys have been reported, and a systematic longitudinal investigation of the added value of such cooperation—compared to exclusively out-of-school offerings—is lacking. More comprehensive collaborations in which the schools provide more than logistical support have yet to be investigated.

Research outside the mentoring field indicates that collaboration between out-of-school and in-school offerings can yield particularly positive effects (Fallik et al., 2013; Itzek-Greulich & Vollmer, 2017; Noam, 2003; Noam & Tillinger, 2004). For example, Itzek-Greulich and Vollmer

(2017) conducted a study with 1,228 secondary school students who were randomly assigned to three different learning environments: a school-only condition, an out-of-school condition (held at a science center), or a combined learning condition for chemistry. Their findings revealed that students in the combined group exhibited the strongest situational competence beliefs. The enhanced effects of the combined learning approach in this and other studies can be partly attributed to differences in organizational, cognitive, affective, and social-environmental aspects between out-of-school and school settings, which can balance and ideally complement one another (Fallik et al., 2013; Sonnett et al., 2025).

Out-of-school mentoring and school offerings differ concerning organizational, cognitive, affective, and social-emotional aspects. For example, organizationally, the relationship between students and mentors in out-of-school mentoring tends to be more symmetrical and non-hierarchical, while the relationship between students and teachers in schools is more hierarchical and governed by stricter norms. The evaluation methods also vary between the two settings. In mentoring, the emphasis is on mutual feedback, whereas in school, formal grades take precedence. From a cognitive standpoint, mentoring often focuses on authentic situations and interdisciplinary knowledge, whereas school offerings prioritize disciplinary knowledge defined by the curriculum. Regarding participants' affective experience, learning in out-of-school contexts is often characterized by excitement, enjoyment, and interest, whereas school learning can frequently evoke negative emotions such as boredom or anxiety (e.g., Vedder-Weiss & Fortus, 2011). Lastly, regarding the social environment, out-of-school activities typically make stronger connections to everyday life and emphasize the needs of individual students, while school-based activities prioritize classroom needs over those of the individual. When planning cooperation between out-of-school activities and school offerings, it is essential to understand the differences in organizational, cognitive, affective, and social-environmental aspects of both settings and how these differences may impact collaboration.

To ensure effective collaboration between out-of-school and school offerings, fundamental bridging principles must be incorporated. Four such bridging principles have been identified as especially relevant (Fallik et al., 2013). First, educational agents in both settings should recognize the importance of cooperation. Second, they must familiarize themselves with the learning content of the other setting and actively seek out content and activities suitable for both settings. Third, it is essential to reduce the novelty space for learners (Neher-Asylbekov & Wagner, 2023; Orion, 1989; Orion & Hofstein, 1994). An adequate preparation of learners involves creating

appropriate prior knowledge for learning activities (especially in an out-of-school context) by introducing new and unfamiliar concepts and familiarizing students with the out-of-school setting, for example, the mentoring setting, as well as preparing them psychologically for the new out-of-school setting and the activities in this setting (e.g., how to approach mentors). Finally, continuous communication and exchange between the participants in the two settings is vital (Noam, 2003; Xplora, 2004).

Factors related to the effectiveness of connecting out-of-school (online) mentoring and school offerings

Research beyond the field of online mentoring indicates that a combination of out-of-school and in-school programs is particularly effective (Fallik et al., 2013; Itzek-Greulich & Vollmer, 2017; Noam, 2003; Noam & Tillinger, 2004). However, the effectiveness of such combinations is likely related to various success factors. In line with the bridging framework by Fallik and colleagues (2013), one factor that should influence the outcome is the extent of bridging between the offerings. The extent of bridging reflects the degree to which learning content, practices, and experiences are meaningfully connected across the two contexts. Strong bridging ensures that insights from the out-of-school learning context are reinforced in school activities and vice versa, reducing novelty and supporting continuity of learning. The more extensive the bridging is, the more successful the combined offerings should be (After School Alliance, 2021; Dahn et al., 2023). Studies outside the field of mentoring support this assumption. For example, Lee and colleagues (2020) investigated the degree of bridging between 279 single-day field-trip programs and school lessons and its effects on various outcomes of 3,721 students from grades 5–8. They found that the degree of bridging was related to the teacher-assessed curriculum-related learning of students and their lesson-related follow-up activities, as well as to student-assessed 21st-century skills and in-curriculum learning. Unfortunately, many studies on in-school–out-of-school collaborations lack detailed descriptions of these partnerships, making systematic analysis difficult (Ziegler et al., 2025b).

Another important factor is the quality of the combined offering. Research outside the field of mentoring supports the importance of quality assessment for program effectiveness. For example, Allen and colleagues (2019) found in a study of 1,599 students in grades 4–12 who participated in 158 different programs, that students who attended programs whose quality had been rated as higher by trained observers had more significant gains in STEM engagement, STEM career interest, STEM career knowledge, and STEM identity than students attending programs with lower levels of program quality.

Finally, the perceived value of STEM among the participants should influence the effectiveness of combined programs. It refers to the social climate surrounding STEM in the peer group. When students perceive that their peers share positive attitudes toward STEM, this creates a supportive group norm that can strengthen individual interest, confidence, and persistence. Numerous studies outside the field of mentoring highlight the importance of this factor. For example, studies show that students who perceived high peer interest in their high school STEM classes are more likely to report STEM career intentions (Hazari et al., 2017). There is also evidence that adolescents' decisions to pursue a STEM major in high school and to enroll in advanced science programs are influenced by their perceptions of peers' interest in STEM (Caspi et al., 2019, 2020). Furthermore, adolescents who perceive peer encouragement and appreciation for STEM report more positive STEM attitudes and more confidence in their STEM abilities (Rice et al., 2013). Finally, aggregated large-scale data at the classroom level show that adolescent girls' self-concept is significantly associated with their female classmates' STEM self-concept and STEM interest (Andersen & Smith, 2024).

Current study

Taken together, research on out-of-school (online) mentoring programs that cooperate with school offerings is scarce. Earlier research has primarily focused on programs where schools handle logistical aspects, such as providing facilities (Finkel, 2017; Guenaga et al., 2022). Additionally, these studies have not employed longitudinal designs or comparative-group designs to evaluate whether such collaborations offer additional participant benefits. However, research outside the field of mentoring has shown that cooperation between out-of-school and school STEM offerings can be particularly effective.

Therefore, it is essential to design out-of-school programs that consider key aspects of effective mentoring for girls in STEM, while also incorporating relevant elements of successful collaborations with schools. Furthermore, empirical research is needed to assess how successful these collaborations are in comparison to standalone out-of-school mentoring programs. Lastly, it is important to investigate critical factors related to the effectiveness of combined programs.

Our goal was to combine an existing out-of-school online mentoring program for girls—based on principles of successful online mentoring and shown to be effective in randomized waitlist control group designs (see Stoeger et al., 2013, 2016, 2023)—with an in-school initiative, namely, teacher-led STEM clubs for mentees. We developed this combined program based on the bridging framework proposed by Fallik et al. (2013), which indicates that the success of integrating out-of-school

and in-school programs depends on aligning organizational, cognitive, affective, and social-environmental factors. Additionally, we aimed to assess the effectiveness of this combined approach. To achieve this, we first examined whether the positive effects of the exclusively out-of-school online mentoring program identified in previous studies could be replicated. Second, we examined whether participation in the combined program has more positive effects than participating in the exclusively out-of-school online mentoring program. Lastly, based on Fallik et al.'s (2013) bridging framework and previous research, we identified three success factors (the extent to which participants perceived a bridging between out-of-school and in-school activities, the quality of the combined program as assessed by the participants, and the participants' perceived value of STEM among the participants) and investigated how they are related to the effectiveness of the combined program.

Research questions

Research Question 1 Is an out-of-school online mentoring program based on key principles of successful mentoring effective in supporting girls in STEM?

To investigate Research Question 1, we selected CyberMentor, a one-year online mentoring program designed for girls, which features female STEM professionals as mentors and includes essential principles of effective mentoring (for more details, see the methods section). In light of previous evaluation studies (Stoeger et al., 2013, 2016), we anticipated that participation would positively impact mentees' elective intentions in STEM, their involvement in STEM activities, their confidence in STEM abilities, and their academic performance in STEM compared to a control group. We used a pretest–posttest control group design to investigate Research Question 1. To ensure that positive effects can be attributed to the out-of-school online mentoring program and not to a special group of students who apply for such programs and thus might generally show more positive developments, mentees' developments were compared to those of a group of girls and boys using weights after implementing propensity score estimation concerning age, school type, and the investigated variables at measurement point 1 (before the mentoring started). While the female control group served as the main reference for evaluating the program's effects on participating girls, the male control group was included to provide an additional benchmark. This enabled us to assess whether the intervention not only promoted development relative to girls who did not participate in the program but also narrowed the typical gender differences observed in STEM domains.

Research Question 2 Does a combined program of out-of-school online mentoring and teacher-led STEM clubs at school that ensures relevant bridging principles lead to more positive effects than exclusively out-of-school online mentoring?

Research indicates that cooperation between school and out-of-school programs yields positive effects (Fallik et al., 2013; Itzek-Greulich & Vollmer, 2017; Noam, 2003; Noam & Tillinger, 2004). Therefore, we anticipated that participation in the combined program would have a more favorable impact on the examined variables than participation in the exclusively out-of-school mentoring program. We used the same design as in Research Question 1 to assess the added value of the combined program compared to the exclusively online mentoring program. To ensure comparability, propensity score estimation was implemented concerning the participants' age, school type, and the investigated variables at measurement point 1 (before the start of the programs).

Research Question 3 Is the effectiveness of the combined program related to the extent of bridging between the out-of-school and school offerings, the quality of the combined program, and the perceived value of STEM among the participants?

Based on the bridging framework by Fallik et al. (2013) and empirical findings outside the mentoring domain, we expected all three variables to be related to the effectiveness of the combined program. For example, research on out-of-school STEM learning shows that students who recognize strong connections between their in-school and out-of-school activities tend to be more engaged and better able to transfer knowledge (Caspi et al., 2019, 2020; Rice et al., 2013). Furthermore, the quality of programs and the value participants place on STEM have been identified as significant predictors of successful outcomes in out-of-school initiatives (Andersen & Smith, 2024; Dahn et al., 2023; Hazari et al., 2017; Lee et al., 2020). Overall, this evidence suggests that bridging perceptions, program quality, and participants' valuing of STEM are critical factors for program effectiveness.

Method

The study setting

CyberMentor as an out-of-school online mentoring program

As an out-of-school online mentoring program, we selected CyberMentor, Germany's most extensive research-based one-on-one online mentoring program for girls and women in STEM. CyberMentor adheres to the key principles of effective mentoring, and its effectiveness has been demonstrated in various studies (Stoeger et al., 2013, 2016, 2017b, 2019, 2021b, 2023). The program aims to spark and maintain girls' interest in STEM; its long-term goal is to increase the participation

rate of girls and women in STEM. Female students in grades 5–13 of secondary education in Germany (mentees) are mentored one-on-one by female out-of-school STEM experts (mentors) who work in STEM or pursue an advanced STEM degree. Mentees participate free of charge, and mentors volunteer their time. Up to 800 mentees and 800 mentors participate in the program annually. Mentors serve as role models, offer insights into their careers and daily work, engage the mentees in discussions about STEM and personal topics, and support mentees with STEM projects. Mentee–mentor dyads are matched based on the mentees' STEM interests and the mentors' STEM fields. Participants commit to a minimum program participation of one year and an interaction frequency of at least 30 min per week. Communication between mentees and mentors, along with collaboration on STEM projects, occurs on a secure online mentoring platform that is exclusively available to members. The project team developed this platform that adheres to all data protection and youth protection regulations. It includes communication tools such as email, text chat, video chat, and forums. Beyond individual mentoring, the platform facilitates networking among all mentors and mentees. It also provides training and support for mentors and mentees through videos, online meetings, and office hours by program staff.

Four quarterly phases structure the mentoring year. In the first phase, the mentees get to know their mentors and learn more about STEM fields and professions. Additionally, the mentees explore and discuss day-to-day STEM phenomena with their mentors. In the second phase, mentees work on STEM projects related to their primary area of interest. Depending on the mentees' age and prior knowledge, the projects can focus on several smaller hands-on projects and experiments or on more extensive, multi-week STEM projects. In the third phase, mentees work on interdisciplinary projects together with their mentors and other mentees and mentors from other STEM fields. The final phase is dedicated to reviewing and reflecting on the mentoring year. Mentees identify and document what they learned and their favorite experiences, for example, by writing articles for a platform-internal STEM magazine edited by the CyberMentor team and participants.

Research has shown the program's short-term, mid-term, and long-term effectiveness (Stoeger et al., 2013, 2016, 2023) and identified important moderating factors, such as the mentees' communication and networking behaviors (Stoeger et al., 2017b, 2021b; Uebler et al., 2023). The research findings have been used to improve the program. The program is endorsed and financed, among others, by the German Federal Ministry of Education and Research and the ministries of education and cultural affairs of various German federal states.

Combining the out-of-school online mentoring program CyberMentor with in-school teacher-led STEM clubs

In an effort to further increase its effectiveness, the out-of-school mentoring program CyberMentor has included a combined version since 2018 that integrates the online program with in-school offerings. Mentees in the combined program participate in the online out-of-school mentoring and additionally attend weekly in-school STEM clubs at their schools. The in-school STEM clubs are run throughout the school year, led by STEM teachers, and consist of students from either a single grade or mixed-grade groups. The number of participants ranges from six to eight mentees per STEM club, ensuring optimal support during practical project work (Kupersmidt et al., 2020). The STEM clubs' activities are guided by the four aforementioned quarterly phases that structure the online mentoring year.

Bridging between the online mentoring program CyberMentor and in-school teacher-led STEM clubs

Program staff offer training programs, online information, and office hours to ensure optimal bridging between out-of-school and in-school offerings. The support focuses on (a) discussions about the importance of cooperation between the educational agents in the two settings, (b) familiarizing teachers with the mentoring setting and supporting them in actively seeking out content and activities suitable for both settings, and (c) helping them to reduce the novelty space for mentees, for example, by creating appropriate knowledge for learning activities, introducing new and unfamiliar concepts, familiarizing students with the out-of-school mentoring setting, and preparing them for the new setting. Furthermore, (d) the program provides an additional online platform with email, text chat, video chat, and forum functions to facilitate regular communication between teachers and mentors.

The STEM teachers who run the STEM clubs are expected to maintain at least quarterly contact with each of their students' mentors to discuss synergies between the out-of-school mentoring and in-school STEM club activities. For example, the teachers inform the mentors about school-based STEM resources (e.g., school labs) and local STEM resources (e.g., nearby science centers or museums). Mentors can then use this information when planning their mentoring sessions. For instance, the STEM club may provide opportunities for students to work hands-on on projects developed during mentoring, and mentors may take local STEM resources into account when discussing STEM activities with their mentees and referring them to relevant offerings. Conversely, the mentors inform the teachers of ways they can support the STEM clubs (e.g., providing data for projects or access to internship opportunities). This collaboration

can benefit the entire group, for example, when the club jointly implements projects suggested by mentors or when mentors present their research during club meetings. Furthermore, the teachers and mentors support the bridging of the mentoring and STEM club activities by encouraging students to bring content from their online mentoring into the STEM club (e.g., project ideas or other STEM topics discussed with their mentors) and vice versa (e.g., discussing STEM club experiments with their mentors or other participants on the online platform).

Sample and procedure

Four groups of students were surveyed in the study: girls who participated in the online mentoring program, girls who participated in the combined program, girls in a control group, and boys in a control group. The members of the control groups participated neither in the online mentoring program nor in the combined program. For the analysis, we only included those girls in the mentoring-only and the combined program groups who had registered for the mentoring program for the first time (as participation for more than one year is possible) and whose mentoring was suitably implemented (i.e., whose mentoring relationship was not terminated prematurely before measurement point 2, who did not experience a mentor change, and who remained active on the online platform at least until measurement point 2, following suggestions by Uebler et al., 2023). Five hundred six participants did not meet these criteria and were excluded from the original data set, resulting in the sample sizes listed in Table 1. Students in the control groups were recruited from schools that implemented the combined program but did not participate in it. They came from 302 different classrooms in 31 public schools in Germany (on average, 9.06 students per classroom participated in the study). We combined data from three school years (2019–2022) to achieve an appropriate sample size for our analyses. All three school years were affected by the COVID-19 pandemic. During the pandemic, schools in

Germany were closed completely for 84 days (OECD, 2021). As with regular schooling, STEM clubs were disrupted during this period. All study participants were enrolled in secondary education, attending grades 5–13. Table 1 contains information on students' age, school type, and response rates at the first and second measurement points for each final sample group.

The data were collected via online questionnaires at two measurement points: at the beginning of the school year (before the programs started) and at the end. Participation in the surveys was voluntary, and participants were assured that their responses would be kept confidential. Written informed consent was obtained from both participants and their legal guardians before the start of the study.

Measures

To evaluate and compare the effectiveness of the two programs, we examined developments in four key areas related to promoting girls in STEM: elective intentions in STEM, STEM activities, STEM grades, and confidence in one's own STEM abilities. We assessed three factors that might influence the effectiveness of the combined program: the extent of bridging between online mentoring in STEM and teacher-led STEM clubs, the quality of the combined program, and the perceived value of STEM among the participants. A list of the items used in the study can be found in the appendix.

Elective intentions in STEM

Participants' elective intentions in STEM were assessed with a five-item scale (Stoeger et al., 2013). Respondents indicated on a 6-point Likert-type scale ranging from 1 (*completely disagree*) to 6 (*completely agree*) how well they could picture themselves choosing a STEM track or course at school, a university STEM major, or a STEM career. A sample item reads: "I can picture myself studying a STEM subject at university." Cronbach's alpha was 0.89 and 0.89 for the two measurement points.

STEM activities

We used a ten-item scale to assess participants' frequency of STEM activities (Stoeger et al., 2021b). Respondents indicated on a Likert-type scale ranging from 1 (*never*) to 6 (*very often*) how frequently they typically engaged in various STEM activities, such as talking about STEM topics with friends or parents, working on a STEM project, or choosing media contributions related to STEM. Cronbach's alpha was 0.88 and 0.97 for the two measurement points.

STEM grades

The STEM grades variable reflects the average of students' self-reported grades from their most recent report

Table 1 Description of study participants and response rates at the two measurement points

Group	<i>n</i>	Age: M (SD)	School type	Re-sponse rate <i>t</i> ₁	Re-sponse rate <i>t</i> ₂
Mentoring only	212	13.66 (2.30)	97%	100%	70%
Combined program	323	12.87 (1.83)	82%	94%	72%
Female control group	1234	13.21 (1.69)	80%	100%	57%
Male control group	1501	12.98 (1.77)	76%	100%	60%

School type denotes the percentage of participants in each group attending college-preparatory secondary schools. In most German federal states, children attend different types of secondary schools as of fifth grade that are either college-preparatory, vocationally focused, or focused on basic skills

card in math, physics, chemistry, biology, and computer science. Grades in Germany range from 1 (*excellent*) to 6 (*insufficient*). For our analysis, we reverse-coded students' grades for the two measurement points so that higher scores indicated higher achievements. Self-reported grades can be assumed to be a reliable indicator of actual achievement (Sticca et al., 2017).

Confidence in STEM abilities

A four-item scale by Dweck and Henderson (1988) was adapted to the STEM domain to assess students' confidence in their abilities in STEM. Each item consists of two statements representing positive and negative self-assessments on opposite ends of a six-point Likert scale. A sample item reads: "I am not very confident about my abilities for the STEM subjects," which contrasts with "I feel fully confident about my abilities for the STEM subjects." Cronbach's alpha was 0.92 and 0.98 for the two measurement points.

Extent of bridging between online mentoring in STEM and teacher-led STEM clubs

The extent of bridging between out-of-school online mentoring and teacher-led STEM clubs at school was measured with three Likert-type items. Participants were asked how often topics from the online mentoring were discussed in the STEM club at school, how often they integrated topics from the STEM club into their online mentoring, and how often they discussed topics from the STEM club on the online mentoring platform. The answer scaled from 1 (*never*) to 6 (*often*). Cronbach's alpha was 0.97.

Quality of the combined program

We assessed how participants perceived the quality of the combined program with six Likert-type items. The answer scaled from 1 (*completely disagree*) to 6 (*completely agree*). An example item reads "Participating in the STEM club was worthwhile." Cronbach's alpha was 0.98.

Perceived value of STEM among the participants

We assessed how participants perceived the STEM attitude of their STEM club peers using three Likert-type items. The answer is scaled from 1 (*completely disagree*) to 6 (*completely agree*). An example item reads, "Most of the girls in my STEM club think STEM subjects are cool." Cronbach's alpha was 0.96.

Handling of missing data

Across all questionnaire variables used in our analyses, 19.97% of values were missing. We conducted Little's (1988) MCAR test to assess whether the data were missing completely at random (MCAR) across the four

student groups. The test indicated that the hypothesis of MCAR must be rejected for the participants in the combined program ($\chi^2(236) = 375, p < 0.001$), the female control group ($\chi^2(137) = 296, p < 0.001$), and the male control group ($\chi^2(202) = 446, p < 0.001$). However, for participants in the mentoring-only program, the test was not significant ($\chi^2(144) = 169, p = 0.078$), suggesting the data to be MCAR.

We conducted multiple imputations to handle missing data as this method is suitable for both data that are MCAR and missing at random (Graham, 2009). The procedure was implemented using the MICE package (version 3.17.0) in R with predictive mean matching for continuous and categorical variables (van Buuren & Groothuis-Oudshoorn, 2006). Following standard guidelines (Graham et al., 2007), we used a large number of imputations ($m = 100$) to ensure stable results. The imputation was performed separately for each of the four groups in our study. For the control groups, the imputation was performed separately for each school year to account for year-specific effects.

Analytical procedure

Initial group comparisons

Based on prior research (Stoeger et al., 2016), which found that participants in the mentoring-only program reported higher elective intentions and STEM activities than the female and the male control groups, we expected differences in the baseline values. Initial group comparisons were conducted using analyses of variance (ANOVA) followed by Tukey's honestly significant difference (HSD) post-hoc tests. Explorative data analyses showed that in three out of four variables, the data of the mentoring-only participants and the data of the participants of the combined program were not normally distributed but rather skewed to the right. These ceiling effects make it difficult to detect increases in outcomes during program participation.

Propensity score weighting

To address these problems and make the four groups comparable, we conducted propensity score estimation to analyze the effectiveness of the out-of-school online mentoring program (Research Question 1) and compare it with the combined program's effectiveness (Research Question 2). Research describes propensity score methods as effective in reducing selection bias in survey data (Stuart, 2010). To adjust for baseline differences between groups, we used generalized boosted models (GBM) with the twang package (version 2.6.1) in R (Cefalu et al., 2024) to estimate propensity scores and weights for the four groups (McCaffrey et al., 2004, 2013). Generalized boosted modeling, a nonparametric tree-based method, effectively captures complex, non-linear relationships

and interactions without relying on assumptions of linearity or additivity among predictors. The propensity score model included baseline values of students' elective intentions in STEM, STEM activities, confidence in STEM abilities, STEM grades, age, and school type as baseline covariates to adjust for potential confounding. Weights for each student were calculated using propensity scores, representing the inverse probability of belonging to the respective group among the four groups, which were then included in the dataset for subsequent weighted analyses for Research Questions 1 and 2 to minimize baseline differences between groups, reducing selection bias and allowing for a more valid estimation of treatment effects when making group comparisons.

Covariates

To reduce potential confounding, we included several covariates in all analyses. First, the respective baseline values of the outcome variables (elective intentions in STEM, STEM activities, confidence in STEM abilities, and STEM grades) were entered as covariates to control for remaining initial differences. Second, for Research Questions 1 and 2, age (measured in years, self-reported by students) was included to account for developmental differences. Finally, for Research Questions 1 and 2, STEM grades (the average of self-reported grades in math, physics, chemistry, biology, and computer science, reverse-coded so that higher values indicated better achievement) were entered as additional covariates in models where they were not the dependent variable.

Analyses of covariances (ANCOVAs)

To evaluate the effectiveness of the out-of-school online mentoring (Research Question 1) and to assess the added value of the combined program (Research Question 2), we compared participants of the mentoring-only and the combined programs with both a female and a male control group, applying propensity score weights to enhance the comparability. We conducted separate one-tailed analyses of covariances (ANCOVAs) for each of the four evaluation variables, using the baseline value of the respective outcome as a covariate to control for initial differences. Doing so, we incorporated propensity score weights with the survey package (version 4.4-2) in R (Lumley, 2024), and results were pooled across multiple imputed datasets using the mitml package (version 0.4-5) in R (Grund et al., 2024). Because our hypotheses were directional and grounded in prior evidence, p -values were derived from one-tailed significance testing. This approach is appropriate when negative effects are neither theoretically expected nor empirically plausible. Covariates were also included to minimize the influence of confounding factors and increase the results' viability. The female control group served as the reference

group, while the male control group was included as an additional dummy variable. To determine whether the combined program led to more positive effects than the exclusively out-of-school online mentoring program (Research Question 2), we compared the effect sizes of the two treatment conditions on the four evaluation variables, following the procedure described by Shrou and Yip-Bannicq (2017).

To assess the relation between the three investigated success factors and the combined program's effectiveness (Research Question 3), one-tailed ANCOVAs were conducted for each of the four evaluation variables, and the three success factors—the extent of bridging between out-of-school mentoring and school offerings, quality of the combined program, and degree of valuing STEM in the participant group—were included as independent variables. Once again, the mitml package (version 0.4-5) in R was used to pool the results across multiple imputed datasets (Grund et al., 2024). The baseline value of the respective evaluation variable was included as a covariate to control for baseline differences. Consistently, one-tailed significance testing was applied due to the directional hypothesis.

Results

Descriptive statistics

Table 2 shows means and standard deviations for each of the investigated evaluation variables at measurement point 1 and measurement point 2, before and after including propensity score estimation weights. Significant baseline differences were observed between the four groups for all investigated variables before performing propensity score estimation. The mentoring-only and the combined program groups had consistently higher means, followed by the male control group. Initial group comparisons using analyses of variance (ANOVA) followed by Tukey's honestly significant difference (HSD) post-hoc tests revealed significant differences ($p < 0.05$) for most baseline values of the investigated evaluation variables across groups. No significant differences emerged in STEM activities between the mentoring-only and the combined program groups ($t(83307) = -1.55, p = 0.122$) and in confidence in STEM abilities between the combined program group and the male control group ($t(21030) = 1.66, p = 0.974$). Evaluating covariate balance across groups showed that incorporating weights calculated through propensity scores reduced these differences substantially, aligning the group means and standard deviations more closely and thus improving the comparability of the groups for subsequent analyses. While some residual baseline differences persisted, the weighting allowed for a more robust estimation of treatment effects. The means and standard deviations for the success factors are as follows: the extent of bridging between

Table 2 Descriptive statistics before and after propensity score estimation

	<i>t</i> ₁		<i>t</i> ₂	
	Before weighting M (SD)	After weighting M (SD)	Before weighting M (SD)	After weighting M (SD)
Elective intentions in STEM				
Mentoring only	4.77 (0.77)	4.04 (1.07)	4.83 (0.88)	4.27 (1.05)
Combined program	4.55 (0.86)	3.93 (1.05)	4.52 (0.98)	4.13 (1.02)
Female control group	3.32 (1.13)	3.67 (1.17)	3.40 (1.18)	3.69 (1.22)
Male control group	3.69 (1.18)	3.69 (1.18)	3.76 (1.18)	3.75 (1.18)
STEM activities				
Mentoring only	3.81 (0.86)	3.73 (0.74)	4.12 (0.81)	4.02 (0.77)
Combined program	3.66 (0.96)	3.47 (0.97)	4.00 (0.81)	3.91 (0.78)
Female control group	3.13 (1.07)	3.33 (1.10)	3.33 (1.05)	3.52 (1.06)
Male control group	3.36 (1.19)	3.33 (1.13)	3.52 (1.11)	3.52 (1.09)
Confidence in STEM abilities				
Mentoring only	4.55 (0.97)	4.28 (0.93)	4.60 (1.05)	4.44 (1.08)
Combined program	4.28 (1.08)	4.05 (1.19)	4.25 (1.14)	4.05 (1.16)
Female control group	3.33 (1.27)	3.82 (1.28)	3.36 (1.34)	3.72 (1.34)
Male control group	4.15 (1.22)	3.95 (1.25)	4.16 (1.17)	4.08 (1.17)
STEM grades				
Mentoring only	5.36 (0.60)	4.97 (0.70)	5.42 (0.52)	5.24 (0.55)
Combined program	5.15 (0.67)	4.94 (0.66)	5.14 (0.69)	4.93 (0.70)
Female control group	4.70 (0.76)	4.82 (0.76)	4.64 (0.79)	4.75 (0.79)
Male control group	4.78 (0.76)	4.82 (0.75)	4.71 (0.78)	4.72 (0.78)

Weighting was performed using propensity score estimation to reduce survey selection bias

online mentoring in STEM and teacher-led STEM club at school, *M* = 3.12, *SD* = 1.21; the quality of the combined program, *M* = 4.79, *SD* = 0.88; and the perceived value of STEM among the participants, *M* = 4.88, *SD* = 0.79.

Research question 1

In Research Question 1, we asked whether an out-of-school online mentoring program based on key principles of successful mentoring is effective in supporting girls in STEM. Table 3 shows the results of the one-tailed ANCOVAs for each of the four evaluation variables, including the baseline value of the respective variable as well as age and STEM grades as covariates. Compared to the female control group as the reference group, participating in the out-of-school online mentoring program significantly positively predicted elective intentions in STEM, STEM activities, confidence in STEM abilities, and STEM grades. For all four evaluation variables, the baseline values of the respective variables also showed significant predictive effects. Age and STEM grades were significant predictors of elective intentions in STEM and confidence in STEM abilities. STEM grades were a significant predictor of STEM activities.

In addition to these positive program effects, an interesting finding emerged. Despite receiving no treatment, the male control group showed significantly greater improvement in confidence in STEM abilities after controlling for pretest scores, compared to the female control group.

The explained variance of the models varied with respect to the four evaluation variables. The model for

Table 3 Results of ANCOVAs for the four evaluation variables for research questions 1 and 2

	Elective intentions in STEM				STEM activities			
	B	SE	β	p	B	SE	β	p
Intercept	-0.13	0.29		0.323	0.95	0.30		<0.001
Baseline score (<i>t</i> ₁)	0.66	0.04	0.65	<0.001	0.52	0.03	0.54	<0.001
Mentoring only	0.31	0.15	0.10	0.023	0.27	0.15	0.10	0.033
Combined program	0.25	0.11	0.09	0.010	0.31	0.09	0.13	<0.001
Male control group	0.04	0.05	0.02	0.204	0.00	0.05	0.00	0.494
Age	0.05	0.02	0.07	0.007	0.02	0.02	0.03	0.117
STEM grades	0.17	0.04	0.10	<0.001	0.12	0.05	0.09	0.012
<i>R</i> ²	0.53				0.38			
	Confidence in STEM abilities				STEM grades			
	B	SE	β	p	B	SE	β	p
Intercept	-0.08	0.40		0.422	1.37	0.20		<0.001
Baseline score (<i>t</i> ₁)	0.55	0.04	0.54	<0.001	0.70	0.03	0.68	<0.001
Mentoring only	0.43	0.24	0.13	0.036	0.38	0.10	0.18	<0.001
Combined program	0.18	0.11	0.06	0.049	0.09	0.05	0.05	0.044
Male control group	0.29	0.05	0.11	<0.001	-0.03	0.03	-0.02	0.175
Age	0.04	0.02	0.06	0.021	0.00	0.01	0.00	0.491
STEM grades	0.23	0.08	0.14	0.002				
<i>R</i> ²	0.42				0.51			

Reference group is the female control group. All *p* values are based on one-tailed testing

elective intentions in STEM explained the most variance at 53%, while the model for STEM activities explained the least at 38%. Thus, our models have strong explanatory power for elective intentions in STEM and STEM grades, moderate explanatory power for confidence in STEM abilities, and relatively lower explanatory power for STEM activities.

In summary, our results show that the out-of-school online mentoring program CyberMentor enhances elective intentions in STEM, STEM activities, confidence in STEM abilities, and STEM grades compared to a propensity-score-weighted control group of girls, after controlling for the investigated variables at measurement point 1, STEM grades, and age.

Research question 2

In Research Question 2, we asked whether a combined program of out-of-school online mentoring and teacher-led STEM clubs at school, that ensures relevant bridging principles, leads to more positive effects than exclusively out-of-school online mentoring.

To assess the added value of combining out-of-school online mentoring with additional in-school support through teacher-led STEM clubs, we conducted one-tailed ANCOVAs for each of the four evaluation variables, as shown in Table 3. Figure 1 shows the development of the means between the two measurement points for all evaluation variables across all groups after weighting. Participation in the combined program also positively affected elective intentions in STEM, STEM activities, confidence in STEM abilities, and STEM grades, compared to the female control group as a reference group. However, contrary to our expectation, one-tailed testing revealed no significant differences among the effect sizes of the combined program and the exclusively out-of-school online mentoring program for predicting students' elective intentions in STEM ($t = -0.28, p = 0.389$), STEM activities ($t = 0.17, p = 0.431$) and confidence in STEM abilities ($t = -0.88, p = 0.191$). Regarding STEM grades, the exclusively out-of-school online mentoring program even had a more positive effect than the combined program ($t = -1.79, p = 0.037$).

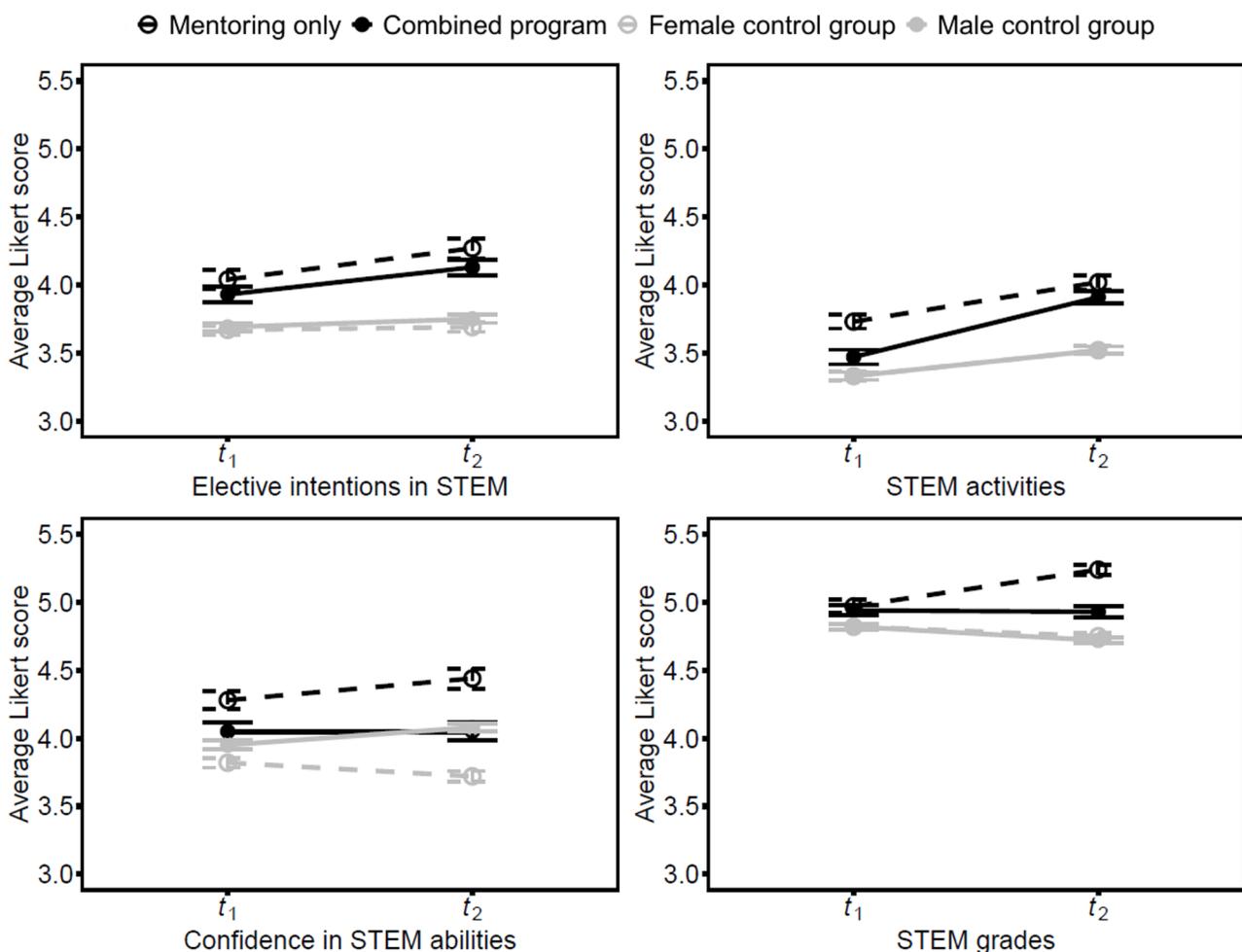


Fig. 1 Longitudinal development of the four groups regarding the four evaluation variables after propensity score estimation

Table 4 Results of ANCOVAs for the four evaluation variables in the combined program for research question 3

	Elective Intentions in STEM				STEM activities			
	B	SE	β	p	B	SE	β	p
Intercept	-0.66	0.42		0.060	0.75	0.39		0.028
Baseline score (t_1)	0.67	0.06	0.59	<0.001	0.41	0.05	0.48	<0.001
Extent of bridging between online mentoring and STEM club	0.05	0.05	0.06	0.181	0.11	0.05	0.16	0.012
Quality of combined program	0.16	0.07	0.15	0.012	0.14	0.08	0.15	0.041
Perceived value of STEM among the participants	0.25	0.08	0.20	0.001	0.16	0.08	0.15	0.026
R^2	0.54				0.40			
	Confidence in STEM abilities				STEM grades			
	B	SE	β	p	B	SE	β	p
Intercept	0.71	0.57		0.110	0.48	0.36		0.091
Baseline score (t_1)	0.53	0.07	0.51	<0.001	0.77	0.05	0.74	<0.001
Extent of bridging between online mentoring and STEM club	-0.02	0.07	-0.03	0.362	0.00	0.03	0.00	0.475
Quality of combined program	0.14	0.11	0.11	0.115	0.08	0.05	0.10	0.052
Perceived value of STEM among the participants	0.14	0.12	0.10	0.126	0.07	0.05	0.08	0.113
R^2	0.33				0.60			

All p values are based on one-tailed testing

In summary, our findings show that while the combined program positively affects elective intentions in STEM, STEM activities, confidence in STEM abilities, and STEM grades, it does not outperform the mentoring-only program.

Research question 3

In Research Question 3, we investigated whether the effectiveness of the combined program is related to the extent of bridging between out-of-school and school offerings, the quality of the combined program, and the perceived value of STEM among the participants.

Table 4 shows the results of ANCOVAs for each of the four evaluation variables, including success factors as independent variables and the baseline values of the respective evaluation variables as covariates. We found relations for elective intentions in STEM and STEM activities, but no relations for confidence in STEM abilities and STEM grades. The effectiveness of the combined program concerning elective intentions in STEM was related to the quality of the combined program and the perceived value of STEM among the participants. The better the students assessed the quality of the combined program and the higher they perceived the value of STEM among STEM club peers, the more positively the combined program influenced their elective intentions in STEM. The effectiveness of the combined program concerning STEM activities was related to all three success factors investigated. The greater the participants assessed the extent of bridging between online mentoring and teacher-led STEM clubs, the more positively they assessed the quality of the combined program, and the more positively they perceived the value of STEM among STEM club peers, the more they increased their STEM activities by participating in the program.

In summary, we found relationships between the investigated success factors and two of the four evaluation variables. The effectiveness of the combined program concerning STEM activities was related to all three investigated success factors (i.e., the extent of bridging between online mentoring in STEM and teacher-led STEM clubs, quality of the combined program, and perceived value of STEM among the participants). The effectiveness of the combined program concerning elective intentions was related to only two of the success factors (i.e., the quality of the combined program and the perceived value of STEM among the participants).

Discussion

The low participation rates of females in STEM fields remain a problematic issue in need of redress (German Federal Statistical Office, 2024; National Center for Education Statistics, 2024). Numerous programs have been initiated over the past few decades to address this challenge. Among these, out-of-school online mentoring programs have shown particular success. These programs begin before the first signs of declining interest in STEM during adolescence and involve women working in STEM professions or pursuing STEM degrees as mentors. They incorporate crucial elements of effective professional mentoring (Luo & Stoeger, 2023; Stoeger et al., 2021a).

The potential for increasing the effectiveness and sustainability of such programs through collaboration with school offerings has not yet been thoroughly explored. This possibility is worthy of investigation, especially given theoretical insights and preliminary research findings from the broader field of STEM education beyond the realm of mentoring (After School Alliance, 2021; Dahn et al., 2023; Fallik et al., 2013; Itzek-Greulich & Vollmer, 2017; Noam, 2003; Noam & Tillinger, 2004). When

investigating how out-of-school–in-school collaborations can improve program outcomes, it is essential to consider the specific bridging principles known to optimize such collaborations (Fallik et al., 2013).

To address this research gap, we developed a combined program that integrates an existing online mentoring program for girls known to be effective (Stoeger et al., 2013, 2016, 2021b, 2023) and in-school teacher-led STEM clubs with weekly meetings. We developed the combined program based on the bridging framework proposed by Fallik et al. (2013), which indicates that the success of integrating out-of-school and in-school programs depends on aligning organizational, cognitive, affective, and social-environmental factors.

In Research Question 1, we explored whether the positive effects identified in previous studies of out-of-school online mentoring on various aspects of girls' STEM development can be replicated. The findings provided a foundation for examining whether a combined program enhances these effects (Research Question 2) and identifying the variables related to its effectiveness (Research Question 3).

As hypothesized for Research Question 1, participation in the out-of-school online mentoring program had a positive effect on all evaluated variables: elective intentions in STEM, engagement in STEM activities, confidence in STEM abilities, and STEM grades. Our results are consistent with previous research demonstrating the positive effects of female mentors, both in programs for college-aged students (Dennehy & Dasgupta, 2017) and in programs for school-aged students (Guenaga et al., 2022). Furthermore, our findings replicated earlier results on the effects of online mentoring for girls aged 10–18 (Stoeger et al., 2013, 2016) while expanding former studies' methodological approach. In their initial survey, Stoeger et al. (2013) compared the development of mentees at three stages: before, during, and after program participation. They employed a waitlist control group consisting of girls who had registered for the program but were randomly assigned to delay their participation by one year. In a subsequent study, Stoeger et al. (2016) compared the development of mentees with both a waitlist control group and control groups of girls and boys who did not register for the program.

In contrast to Stoeger et al. (2016), we compared the development of mentees with that of girls and boys who did not participate in the program but were comparable to program participants by implementing propensity score estimation weights regarding age, school type, and baseline values of all evaluation variables at the first measurement point. This approach imposed a stricter criterion than that used by Stoeger et al. (2016), ensuring that the control group of participants who did not sign up for the program was as comparable as possible to

the participants. This is particularly important because research indicates that students who enroll in long-term programs such as CyberMentor often have significantly more favorable baseline characteristics, such as higher interest in STEM, greater confidence in their abilities, and better grades in STEM subjects, compared to those who do not enroll (Stoeger et al., 2016). In our sample, the girls who applied to the online mentoring program also exhibited significantly higher baseline values across all evaluation variables. However, we ensured the groups were as comparable as possible by using propensity score estimation and controlling for baseline values in our analyses. This allowed us to more confidently attribute the observed effects to the mentoring program rather than to inherent advantages among the students who chose to register for such programs and who likely would have shown positive developments.

It is noteworthy that, despite receiving no intervention, the male control group exhibited a significantly greater improvement in confidence in STEM abilities after controlling for pretest scores than the female control group. Gender differences in STEM confidence of children and youths are well-documented in previous research, with boys typically showing higher confidence than girls (Else-Quest et al., 2010; Seo et al., 2019). One potential explanation is that prevailing stereotypes regarding boys' and girls' abilities shape children's and youths' beliefs in their own STEM competencies (Master, 2021). As stereotypes about STEM typically favor boys (Bian et al., 2017; Miller et al., 2018), our findings may reflect a stereotype-boost effect on boys' confidence in their STEM abilities, as shown in previous research (Kim et al., 2023; Kurtz-Costes et al., 2008).

In Research Question 2, we investigated whether combining out-of-school online mentoring with weekly teacher-led STEM clubs—especially when considering the four bridging principles outlined by Fallik et al. (2013)—yields more significant effects than online mentoring alone. To our knowledge, this question has not yet been explored. Furthermore, only two other out-of-school mentoring programs have been designed to collaborate with schools on supporting girls in STEM (Finkel, 2017; Guenaga et al., 2022). However, neither study assessed whether this collaboration with schools provided additional benefits compared to out-of-school mentoring alone. Moreover, in both instances, the school collaborations were strictly logistical. For example, the schools helped address logistical challenges, such as finding meeting spaces, ensuring laboratory safety, and securing on-site materials. Other than the program we described in this study, we know of no mentoring programs in which cooperation with schools extends beyond logistical aspects. Insights from the limited existing studies on STEM education outside the mentoring context, as

well as theoretical considerations, make it reasonable to assume that such programs might lead to more positive and sustainable outcomes than traditional extracurricular mentoring—especially when relevant bridging principles are considered (Fallik et al., 2013; Itzek-Greulich & Vollmer, 2017; Noam, 2003; Noam & Tillinger, 2004; Sonnert et al., 2025).

Contrary to our assumption, a combination of out-of-school online mentoring and in-school teacher-led STEM clubs was not generally more effective than the out-of-school online mentoring program alone. Although participation in the combined program had positive effects on all evaluation variables (elective intentions in STEM, STEM activities, confidence in STEM abilities, and STEM grades), these effects were comparable to those of the mentees in the exclusively out-of-school online mentoring program.

As noted above, to our knowledge, our study is the first to examine the added benefit of combining in-school support with out-of-school mentoring. However, our results can be compared with previous research on collaborations between in-school and out-of-school STEM programs outside the mentoring field. For example, Itzek-Greulich and Vollmer (2017) reported positive effects of both an out-of-school lab work condition and a combined lab work condition relative to a no-treatment condition, consistent with our findings. However, diverging from our results, the combined condition yielded stronger effects on feelings of competence than the purely out-of-school condition.

Our findings can be better understood when viewed through the lens of the bridging framework (Fallik et al., 2013). As predicted by theory and consistent with the results of Research Question 3, the combined program did not generally outperform mentoring alone; its effectiveness was related to participating students' perceptions of both the extent of bridging between online STEM mentoring and teacher-led STEM clubs and the quality of the combined program. These results are consistent with findings from studies outside the field of mentoring, which indicated that more favorable STEM outcomes were associated with the degree of bridging between out-of-school offerings and school lessons (Lee et al., 2020; Neher-Asylbekov & Wagner, 2023) as well as the quality of STEM programs (Allen et al., 2019). This suggests that any potential added value of the combined program would likely depend on the optimal implementation of bridging and the quality of the combined offerings. This should apply not only to the implementation of the STEM clubs and the mentoring program, but also to the quality of the quarterly interactions between teachers and mentors. Both aspects warrant further investigation in future research.

The effectiveness of the combined program was also related to participants' perceived value of STEM among the STEM club participants. This finding aligns with prior research showing associations between the peer group climate regarding STEM and individual STEM outcomes, such as career intentions and self-confidence (Andersen & Smith, 2024; Caspi et al., 2019, 2020; Hazari et al., 2017). Consistent with this research, our results underscore the importance of fostering a positive peer environment around STEM.

A further reason the combined program was not generally more effective than the pure online mentoring program may be related to the implementation challenges posed by the COVID-19 pandemic. During the pandemic, secondary schools in Germany were closed for 84 days, disrupting regular school activities and STEM clubs. As a result, there were fewer teacher-led STEM clubs than initially planned, which might have affected the perceived bridging, quality, and valuing of STEM among participants, ultimately leading to an underestimation of the added value of combining the online mentoring program with in-school STEM activities.

Another reason the combined program did not generally outperform the online mentoring program might be that participants' baseline values on the evaluation variables are quite high. Due to ceiling effects, finding positive effects of the exclusively out-of-school online mentoring program is already challenging. Consequently, expecting an additional effect from the combined program becomes even more improbable.

Limitations and future work

Our study is the first to investigate whether a combination of out-of-school online mentoring and in-school teacher-led STEM clubs has a more positive impact on advancing girls in STEM than a proven-effective out-of-school online mentoring program. The findings provide valuable insights into the advancement of girls in STEM fields. However, it is important to note that our study has some limitations.

First, as described above, the implementation of the combined program was not optimal due to the reduced number of weekly teacher-led STEM clubs during the COVID-related school closures and other pandemic-related difficulties. In the future, it would, therefore, be important to implement the combined program under non-COVID conditions and investigate if the effects might be even stronger.

A second limitation was that we measured the success factors (extent of bridging between online mentoring in STEM and teacher-led STEM clubs, quality of the combined program, and perceived value of STEM among the participants) of the combined program via student assessments. In future studies, the implementation of

bridging and the quality of the combined offerings should be assessed with more objective measures. Furthermore, the quality and specific content of the quarterly interactions between teachers and mentors should be examined. For this purpose, teachers and mentors could fill out diaries and checklists. Classroom observations could also be conducted, enabling an even more objective assessment. Finally, it would be desirable to conduct interviews on the conditions for successful bridging with teachers and mentors.

Program outcomes were also measured exclusively by student assessments, which can be considered a third limitation. Future research may benefit from broadening the focus to include the perspectives of teachers and mentors, for example, by incorporating teacher and mentor reports on program implementation and student outcomes.

A fourth limitation concerns the generalizability of our findings to other STEM programs, age groups, and countries. Programs such as CyberMentor often demand substantial time, financial resources, and robust legal safeguards to meet strict data protection and child protection requirements. Nevertheless, we believe our results offer valuable insights that can inform the design and implementation of similar programs for young people. Although the program investigated in our study represents a specific support context, our findings align well with previous research both within and beyond the field of mentoring. This suggests that they may inform future program development not only for other mentoring initiatives seeking to collaborate with schools but also for in-school and out-of-school collaborations more generally.

Consistent with our empirical findings, which highlight the central role of bridging in the success of such collaborations, feedback from participating teachers and mentors emphasized the reciprocal synergies that emerged through the collaboration as particularly beneficial. At the same time, both groups described additional time demands as challenging and, in some cases, noted a perceived lack of individual engagement from their respective counterparts – indicating that effective bridging was not always achieved. These observations are based on informal feedback only, underscoring the need for further research on the benefits and challenges of such collaborations from the perspectives of both in-school and out-of-school stakeholders.

Conclusion and practical implications

Our study showed for the first time that combining an out-of-school online mentoring program for girls in STEM with in-school teacher-led STEM clubs effectively influences elective intentions in STEM, STEM activities, confidence in STEM abilities, and STEM grades among

girls. The combined program was particularly effective when participants perceived a high level of bridging between out-of-school and school offerings and rated the quality of the combined program and the appreciation of STEM within the participating peer group highly. In practice, this means that the combination of out-of-school online mentoring and school-based support leads to positive effects for girls, however, its effectiveness is contingent upon adequate implementation through extensive bridging.

The findings of this study imply several ways to enhance the effectiveness of (online) mentoring programs in STEM. Program designers should not only attend to key issues in professionalizing mentoring practices (Stoeger et al., 2021a), such as tailoring programs to their specific context and participants, clarifying and aligning goals and expectations of mentors and mentees, and ensuring adequate resources. They should also find ways to systematically connect out-of-school mentoring with in-school activities, such as providing teacher-led STEM clubs at school. When doing so, it is essential to attend to relevant bridging principles (Fallik et al., 2013). For example, there should be structured opportunities for mentors and teachers to exchange information, to jointly plan projects and discuss how to implement them, and to collaboratively support mentees in linking their experiences across different contexts. Throughout the process, it is essential to maintain a strong focus on the quality of the combined program. In addition to the strategies noted above for promoting effective bridging, practical steps to enhance quality include ensuring that both the mentoring platform and school-based components are coherent, engaging, and well supported through adequate training and resources for mentors and teachers, as well as ongoing monitoring. Formative evaluations can help to adapt and improve both components continuously. Lastly, fostering a positive peer environment around STEM is crucial. Mentors and teachers can promote collaborative group projects, recognize peer achievements, and facilitate discussions on the importance of STEM in everyday life. By intentionally addressing these dimensions, future online mentoring programs can significantly enhance girls' engagement in STEM.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40594-026-00600-7>.

Supplementary Material 1

Acknowledgements

We thank Marina Kaas and Michael Heilemann for the technical implementation of the surveys, Benjamin Matthes for his assistance in conducting the survey, and Daniel Patrick Balestrini for language feedback.

Author contributions

H.S. conceptualized the study together with A.Z. and played the lead role in writing the manuscript. C.U. coordinated the survey. C.U. and S.B. conducted the data preprocessing and the statistical analysis and were major contributors to writing the methodology and results sections. K.E. contributed to the literature review and the theory, methods and discussion section. All authors were involved in the critical review, and the final manuscript was approved.

Funding

Open Access funding enabled and organized by Projekt DEAL. This research was funded by the German Federal Ministry of Education and Research (BMBF) under grant number 01JW1801N and by the Bavarian State Ministry for Education and Culture. The responsibility for all content supplied lies with the authors.

Data availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request. The data are not publicly available due to local data protection regulations.

Declarations

Human ethics and consent to participate

Participation in the surveys was voluntary. Written informed consent was obtained from both participants and their legal guardians before the start of the study.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Educational Sciences, University of Regensburg, Regensburg, Germany

²Department of Educational Psychology, University of Erlangen-Nuremberg, Erlangen, Germany

Received: 25 August 2025 / Accepted: 28 January 2026

Published online: 25 February 2026

References

- After School Alliance (2021). *STEM learning in afterschool is on the rise, but barriers and inequities exist*. <http://afterschoolalliance.org/documents/AA3PM/AA3PM-STEM-Report-2021.pdf>
- Allen, P. J., Chang, R., Gorrall, B. K., Waggenspack, L., Fukuda, E., Little, T. D., & Noam, G. G. (2019). From quality to outcomes: A National study of afterschool STEM programming. *International Journal of STEM Education*, 6(1), 37. <https://doi.org/10.1186/s40594-019-0191-2>
- Andersen, I. G., & Smith, E. (2024). Gender differences in math and science academic self-concepts and the association with female climate in 8th grade classrooms. *The Journal of Early Adolescence*, 44(5), 545–578. <https://doi.org/10.1177/02724316231188682>
- Bian, L., Leslie, S. J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, 355(6323), 389–391. <https://doi.org/10.1126/science.aah6524>
- Blake-Beard, S., Bayne, M. L., Crosby, F. J., & Muller, C. B. (2011). Matching by race and gender in mentoring relationships: Keeping our eyes on the prize. *Journal of Social Issues*, 67(3), 622–643. <https://doi.org/10.1111/j.1540-4560.2011.01717.x>
- Canaan, S., & Mouganie, P. (2023). The impact of advisor gender on female students' STEM enrollment and persistence. *Journal of Human Resources*, 58(2), 593–632. <https://doi.org/10.3368/jhr.58.4.0320-10796R2>
- Caspi, A., Gorsky, P., Nitzani-Hendel, R., Zacharia, Z., Rosenfeld, S., Berman, S., & Shildhouse, B. (2019). Ninth-grade students' perceptions of the factors that led them to major in high school science, technology, engineering, and mathematics disciplines. *Science Education*, 103(5), 1176–1205. <https://doi.org/10.1002/sce.21524>
- Caspi, A., Gorsky, P., Nitzani-Hendel, R., Zacharia, Z. C., Rosenfeld, S., Berman, S., & Shildhouse, B. (2020). Children's perceptions of the factors that led to their enrolment in advanced, middle-school science programmes. *International Journal of Science Education*, 42(11), 1915–1939. <https://doi.org/10.1080/09500693.2020.1802083>
- Cefalu, M., Ridgeway, G., McCaffrey, D., Morral, A., Ann Griffin, B., & Burgette, L. (2024). Twang: Toolkit for weighting and analysis of nonequivalent groups. *R Package Version 2.6.1*. <https://doi.org/10.32614/CRAN.package.twang>
- Dahn, M., Peppler, K., & Ito, M. (2023). Making connections to and from out-of-school experiences. *Review of Research in Education*, 47(1), 443–473. <https://doi.org/10.3102/0091732X231211271>
- Dasgupta, N. (2011). Ingroup experts and peers as social vaccines who inoculate the self-concept: The stereotype inoculation model. *Psychological Inquiry*, 22(4), 231–246. <https://doi.org/10.1080/1047840X.2011.607313>
- Dennehy, T. C., & Dasgupta, N. (2017). Female peer mentors early in college increase women's positive academic experiences and retention in engineering. *Proceedings of the National Academy of Sciences of the United States of America*, 114(23), 5964–5969. <https://doi.org/10.1073/pnas.1613117114>
- Dick, T. J. M., Besomi, M., Coltman, C. E., Diamond, L. E., Hall, M., Maharaj, J., Kean, C. O., Barzan, M., & Mickle, K. J. (2024). A 100-day mentoring program leads to positive shifts in girls' perceptions and attitudes towards biomechanics and related STEM disciplines. *Journal of Biomechanics*, 173, 112244. <https://doi.org/10.1016/j.jbiomech.2024.112244>
- DuBois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. *American Journal of Community Psychology*, 30(2), 157–197. <https://doi.org/10.1023/a:1014628810714>
- DuBois, D. L., Portillo, N., Rhodes, J. E., Silverthorn, N., & Valentine, J. C. (2011). How effective are mentoring programs for youth? A systematic assessment of the evidence. *Psychological Science in the Public Interest*, 12(2), 57–91. <https://doi.org/10.1177/1529100611414806>
- Dweck, C. S., & Henderson, V. L. (1988). *Theories of intelligence: Background and measures* [Unpublished manuscript]. University of Illinois. <https://eric.ed.gov/?id=ED312057>
- Eby, L. T., Allen, T. D., Evans, S. C., Ng, T., & DuBois, D. L. (2008). Does mentoring matter? A multidisciplinary meta-analysis comparing mentored and non-mentored individuals. *Journal of Vocational Behavior*, 72(2), 254–267. <https://doi.org/10.1016/j.jvb.2007.04.005>
- Eby, L. T., Allen, T. D., Hoffman, B. J., Baranik, L. E., Sauer, J. B., Baldwin, S., Morrison, M. A., Kinkade, K. M., Maher, C. P., Curtis, S., & Evans, S. C. (2013). An interdisciplinary meta-analysis of the potential antecedents, correlates, and consequences of protégé perceptions of mentoring. *Psychological Bulletin*, 139(2), 441–476. <https://doi.org/10.1037/a0029279>
- Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, 136(1), 103–127. <https://doi.org/10.1037/a0018053>
- Emmerding, K., Heyna, K., & Stöger, H. (2024). E-Mentoring: Ein vielversprechendes Format für die Interessens- und Talentförderung junger Menschen [E-mentoring: A promising method for promoting the interests and talents of young people]. In G. Jakob, & B. Schüler (Eds.), *Patenschaften und Mentoring für Kinder und Jugendliche. Eine neue Kultur des Engagements zur Förderung von Bildung, Teilhabe und Integration* (pp. 128–136). Beltz Juventa.
- Fallik, O., Rosenfeld, S., & Eylon, B. S. (2013). School and out-of-school science: A model for bridging the gap. *Studies in Science Education*, 49(1), 69–91. <https://doi.org/10.1080/03057267.2013.822166>
- Finkel, L. (2017). Walking the path together from high school to STEM majors and careers: Utilizing community engagement and a focus on teaching to increase opportunities for URM students. *Journal of Science Education and Technology*, 26(1), 116–126. <https://doi.org/10.1007/s10956-016-9656-y>
- Garringer, M., Kupersmidt, J. B., Rhodes, J., Stelter, R., & Tai, T. (2015). *Elements of effective practice for mentoring: Research-informed and practitioner-approved best practices for creating and sustaining impactful mentoring relationships and strong program services* (4th edition). MENTOR. <https://files.eric.ed.gov/fulltext/ED594036.pdf>
- German Federal Statistical Office (2024). *Students enrolled in STEM courses*. <https://www.destatis.de/EN/Themes/Society-Environment/Education-Research-Culture/Institutions-Higher-Education/Tables/students-in-stem-courses.html>
- Gladstone, J. R., & Cimpian, A. (2021). Which role models are effective for which students? A systematic review and four recommendations for maximizing the effectiveness of role models in STEM. *International Journal of STEM Education*, 8(1), 59. <https://doi.org/10.1186/s40594-021-00315-x>
- González-Pérez, S., de Cabo, M. R., & Sáinz, M. (2020). Girls in STEM: Is it a female role-model thing? *Frontiers in Psychology*, 11, 2204. <https://doi.org/10.3389/fpsyg.2020.02204>

- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60, 549–576. <https://doi.org/10.1146/annurev.psych.58.110405.085530>
- Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science*, 8(3), 206–213. <https://doi.org/10.1007/s1121-007-0070-9>
- Grossman, J. B., & Rhodes, J. E. (2002). The test of time: Predictors and effects of duration in youth mentoring relationships. *American Journal of Community Psychology*, 30(2), 199–219. <https://doi.org/10.1023/A:1014680827552>
- Grossman, J. B., Chan, C. S., Schwartz, S. E. O., & Rhodes, J. E. (2012). The test of time in school-based mentoring: The role of relationship duration and re-matching on academic outcomes. *American Journal of Community Psychology*, 49(1–2), 43–54. <https://doi.org/10.1007/s10464-011-9435-0>
- Grund, S., Robitzsch, A., & Lüdtke, O. (2024). *Mitml: Tools for multiple imputation in multilevel modeling*. R Package Version 0.4–5.
- Guenaga, M., Eguiluz, A., Garaizar, P., & Mimenza, A. (2022). The impact of female role models leading a group mentoring program to promote STEM vocations among young girls. *Sustainability*, 14(3), 1420. <https://doi.org/10.3390/su14031420>
- Hazari, Z., Potvin, G., Cribbs, J. D., Godwin, A., Scott, T. D., & Klotz, L. (2017). Interest in STEM is contagious for students in biology, chemistry, and physics classes. *Science Advances*, 3(8), e1700046. <https://doi.org/10.1126/sciadv.1700046>
- Hernandez, P. R., Bloodhart, B., Barnes, R. T., Adams, A. S., Clinton, S. M., Pollack, I., Godfrey, E., Burt, M., & Fischer, E. V. (2017). Promoting professional identity, motivation, and persistence: Benefits of an informal mentoring program for female undergraduate students. *PLoS One*, 12(11), e0187531. <https://doi.org/10.1371/journal.pone.0187531>
- Hopp, M. D. S., Stoeger, H., & Ziegler, A. (2020). The supporting role of mentees' peers in online mentoring: A longitudinal social network analysis of peer influence. *Frontiers in Psychology*, 11, 1929. <https://doi.org/10.3389/fpsyg.2020.01929>
- Ito, M., Gutiérrez, K., Livingstone, S., Penuel, B., Rhodes, J., Salen, K., Schor, J., Sefton-Green, J., & Watkins, S. C. (2013). *Connected learning: An agenda for research and design*. Digital Media and Learning Research Hub. <https://eprints.lse.ac.uk/48114>
- Ito, M., Arum, R., Conely, D., Gutiérrez, K., Krischner, B., Livingstone, S., Michalchik, V., Pepler, K., Penuel, W., Pinkard, N., Rhodes, J., Salen Tekinbaş, K., & Schor, J. (2020). *The connected learning research network: Reflections on a decade of engaged scholarship*. Deakin University. https://dro.deakin.edu.au/articles/report/the_connected_learning_research_network_reflections_on_a_decade_of_engaged_scholarship/20619648
- Itzek-Greulich, H., & Vollmer, C. (2017). Emotional and motivational outcomes of lab work in the secondary intermediate track: The contribution of a science center outreach lab. *Journal of Research in Science Teaching*, 54(1), 3–28. <https://doi.org/10.1002/tea.21334>
- Jethwani, M. M., Memon, N., Seo, W., & Richer, A. (2017). I can actually be a super sleuth: Promising practices for engaging adolescent girls in cybersecurity education. *Journal of Educational Computing Research*, 55(1), 3–25. <https://doi.org/10.1177/0735633116651971>
- Keller, T. E., & DuBois, D. L. (2021). Influence of program staff on quality of relationships in a community-based youth mentoring program. *Annals of the New York Academy of Sciences*, 1483(1), 112–126. <https://doi.org/10.1111/nyas.14289>
- Kim, J., Liu, R., & Zhao, X. (2023). A big (male) fish in a small pond? The gendered effect of relative ability on STEM aspirations under stereotype threat. *European Sociological Review*, 39(2), 177–193. <https://doi.org/10.1093/esr/jcac037>
- Kupersmidt, J. B., Stelter, R., Garringer, M., & Bourgoin, J. (2018). *STEM Mentoring: Supplement to the elements of effective practice for mentoring*. MENTOR. <https://www.mentoring.org/resource/stem-mentoring-supplement-to-the-element-s-of-effective-practice-for-mentoring/>
- Kupersmidt, J. B., Stelter, R. L., Kuperminc, G., Garringer, M., & Shane, J. (2020). *Group mentoring: Supplement to the elements of effective practice for mentoring*. MENTOR. https://www.mentoring.org/wp-content/uploads/2020/09/group-mentoring-supplement-to-EEP_final.pdf
- Kurtz-Costes, B., Rowley, S. J., Harris-Britt, A., & Woods, T. A. (2008). Gender stereotypes about mathematics and science and self-perceptions of ability in late childhood and early adolescence. *Merrill-Palmer Quarterly*, 54(3), 386–409. <https://doi.org/10.1353/mpq.0.0001>
- Lawlor, G., Byrne, P., & Tangney, B. (2020). CodePlus: Measuring short-term efficacy in a non-formal, all-female CS outreach programme. *ACM Transactions on Computing Education*, 20(4), 1–18. <https://doi.org/10.1145/3411510>
- Lawner, E. K., Quinn, D. M., Camacho, G., Johnson, B. T., & Pan-Weisz, B. (2019). Ingroup role models and underrepresented students' performance and interest in STEM: A meta-analysis of lab and field studies. *Social Psychology of Education*, 22(5), 1169–1195. <https://doi.org/10.1007/s11218-019-09518-1>
- Lee, H., Stern, M. J., & Powell, R. B. (2020). Do pre-visit preparation and post-visit activities improve student outcomes on field trips? *Environmental Education Research*, 26(7), 989–1007. <https://doi.org/10.1080/13504622.2020.1765991>
- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), 1198. <https://doi.org/10.2307/2290157>
- Lumley, T. (2024). *Survey: Analysis of complex survey samples*. R Package Version 4.4–2.
- Luo, L., & Stoeger, H. (2023). Unlocking the transformative power of mentoring for youth development in communities, schools, and talent domains. *Journal of Community Psychology*, 51(8), 3067–3082. <https://doi.org/10.1002/jcop.23082>
- Lyons, M. D., & Edwards, K. D. (2022). Strategies for monitoring mentoring relationship quality to predict early program dropout. *American Journal of Community Psychology*, 70(1–2), 127–138. <https://doi.org/10.1002/ajcp.12585>
- Martín-Peciña, M., Quesada, A., Abril, A. M., & Romero-Ariza, M. (2025). Breaking barriers to unleash STEM futures by empowering girls through mentorship in summer camps. *Education Sciences*, 15(2), 242. <https://doi.org/10.3390/educsci15020242>
- Master, A. (2021). Gender stereotypes influence children's STEM motivation. *Child Development Perspectives*, 15(3), 203–210. <https://doi.org/10.1111/cdep.12424>
- McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score Estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods*, 9(4), 403–425. <https://doi.org/10.1037/1082-989X.9.4.403>
- McCaffrey, D. F., Griffin, B. A., Almirall, D., Slaughter, M. E., Ramchand, R., & Burgette, L. F. (2013). A tutorial on propensity score Estimation for multiple treatments using generalized boosted models. *Statistics in Medicine*, 32(19), 3388–3414. <https://doi.org/10.1002/sim.5753>
- McQuillin, S. D., & Lyons, M. D. (2021). A National study of mentoring program characteristics and premature match closure: The role of program training and ongoing support. *Prevention Science*, 22(3), 334–344. <https://doi.org/10.1007/s1121-020-01200-9>
- Miller, D. I., Nolla, K. M., Eagly, A. H., & Uttal, D. H. (2018). The development of children's gender-science stereotypes: A meta-analysis of 5 decades of U.S. Draw-a-scientist studies. *Child Development*, 89(6), 1943–1955. <https://doi.org/10.1111/cdev.13039>
- National Center for Education Statistics (2024). *Science, technology, engineering, and mathematics (STEM) education, by gender*. <https://nces.ed.gov/fastfacts/display.asp?id=899>
- Neher-Asylbekov, S., & Wagner, I. (2023). Effects of out-of-school STEM learning environments on student interest: A critical systematic literature review. *Journal for STEM Education Research*, 6(1), 1–44. <https://doi.org/10.1007/s41979-022-00080-8>
- Noam, G. G. (2003). Learning with excitement: Bridging school and after-school worlds and project-based learning. *New Directions for Youth Development*, 2003(97), 121–138. <https://doi.org/10.1002/ym.39>
- Noam, G. G., & Tillinger, J. R. (2004). After-school as intermediary space: Theory and typology of partnerships. *New Directions for Youth Development*, 2004(101), 75–113. <https://doi.org/10.1002/ym.73>
- OECD. (2021). The state of global education: 18 months into the pandemic. *OECD Publishing*. <https://doi.org/10.1787/1a23bb23-en>
- Orion, N. (1989). Development of a high-school geology course based on field trips. *Journal of Geological Education*, 37(1), 13–17. <https://doi.org/10.5408/0022-1368-37.1.13>
- Orion, N., & Hofstein, A. (1994). Factors that influence learning during a scientific field trip in a natural environment. *Journal of Research in Science Teaching*, 31(10), 1097–1119. <https://doi.org/10.1002/tea.3660311005>
- Prieto-Rodríguez, E., Sincoc, K., & Blackmore, K. (2020). STEM initiatives matter: Results from a systematic review of secondary school interventions for girls. *International Journal of Science Education*, 42(7), 1144–1161. <https://doi.org/10.1080/09500693.2020.1749909>
- Raposa, E. B., Rhodes, J., Stams, G. J. J. M., Card, N., Burton, S., Schwartz, S., Sykes, L. A. Y., Kanchewa, S., Kupersmidt, J. B., & Hussain, S. (2019). The effects of youth mentoring programs: A meta-analysis of outcome studies. *Journal of Youth and Adolescence*, 48(3), 423–443. <https://doi.org/10.1007/s10964-019-00982-8>
- Rice, L., Barth, J. M., Guadagno, R. E., Smith, G. P. A., & McCallum, D. M. (2013). The role of social support in students' perceived abilities and attitudes toward math and science. *Journal of Youth and Adolescence*, 42(7), 1028–1040. <https://doi.org/10.1007/s10964-012-9801-8>

- Sáinz, M., Fàbregues, S., Romano, M. J., & López, B. S. (2022). Interventions to increase young people's interest in STEM: A scoping review. *Frontiers in Psychology, 13*, 954996. <https://doi.org/10.3389/fpsyg.2022.954996>
- Santo, R., Ching, D., Hoadley, C., & Peppler, K. (2019). *Brokering youth pathways: A toolkit for connecting youth to future opportunity*. Hive Research Lab. https://brokering.hiveresearchlab.org/wp-content/uploads/2019/01/hrl_2018_brokerimg_full_brief_set.pdf
- Sasson, I. (2019). Participation in research apprenticeship program: Issues related to career choice in STEM. *International Journal of Science and Mathematics Education, 17*(3), 467–482. <https://doi.org/10.1007/s10763-017-9873-8>
- Seo, E., Shen, Y., & Alfaro, E. C. (2019). Adolescents' beliefs about math ability and their relations to STEM career attainment: Joint consideration of race/ethnicity and gender. *Journal of Youth and Adolescence, 48*(2), 306–325. <https://doi.org/10.1007/s10964-018-0911-9>
- Shrout, P. E., & Yip-Bannicq, M. (2017). Inferences about competing measures based on patterns of binary significance tests are questionable. *Psychological Methods, 22*(1), 84–93. <https://doi.org/10.1037/met0000109>
- Sonnert, G., Reid, T., Sunbury, S., & Sadler, P. M. (2025). How do formal and informal science learning experiences during high school shape students' career interest and STEM identity? *International Journal of STEM Education, 12*(1), 55. <https://doi.org/10.1186/s40594-025-00568-w>
- Stelter, R. L., Kupersmidt, J. B., & Stump, K. N. (2021). Establishing effective STEM mentoring relationships through mentor training. *Annals of the New York Academy of Sciences, 1483*(1), 224–243. <https://doi.org/10.1111/nyas.14470>
- Sticca, F., Goetz, T., Bieg, M., Hall, N. C., Eberle, F., & Haag, L. (2017). Examining the accuracy of students' self-reported academic grades from a correlational and a discrepancy perspective: Evidence from a longitudinal study. *PLoS One, 12*(11), e0187367. <https://doi.org/10.1371/journal.pone.0187367>
- Stoeger, H., Duan, X., Schirner, S., Greindl, T., & Ziegler, A. (2013). The effectiveness of a one-year online mentoring program for girls in STEM. *Computers & Education, 69*, 408–418. <https://doi.org/10.1016/j.compedu.2013.07.032>
- Stoeger, H., Schirner, S., Laemmle, L., Obergriesser, S., Heilemann, M., & Ziegler, A. (2016). A contextual perspective on talented female participants and their development in extracurricular STEM programs. *Annals of the New York Academy of Sciences, 1377*(1), 53–66. <https://doi.org/10.1111/nyas.13116>
- Stoeger, H., Hopp, M., & Ziegler, A. (2017b). Online mentoring as an extracurricular measure to encourage talented girls in STEM (science, technology, engineering, and mathematics): An empirical study of one-on-one versus group mentoring. *Gifted Child Quarterly, 61*(3), 239–249. <https://doi.org/10.1177/0016986217702215>
- Stoeger, H., Debatin, T., Heilemann, M., & Ziegler, A. (2019). Online mentoring for talented girls in STEM: The role of relationship quality and changes in learning environments in explaining mentoring success. *New Directions for Child and Adolescent Development, 2019*(168), 75–99. <https://doi.org/10.1002/cad.20320>
- Stoeger, H., Balestrini, D. P., & Ziegler, A. (2021a). Key issues in professionalizing mentoring practices. *Annals of the New York Academy of Sciences, 1483*(1), 5–18. <https://doi.org/10.1111/nyas.14537>
- Stoeger, H., Heilemann, M., Debatin, T., Hopp, M. D. S., Schirner, S., & Ziegler, A. (2021b). Nine years of online mentoring for secondary school girls in STEM: An empirical comparison of three mentoring formats. *Annals of the New York Academy of Sciences, 1483*(1), 153–173. <https://doi.org/10.1111/nyas.14476>
- Stoeger, H., Debatin, T., Heilemann, M., Schirner, S., & Ziegler, A. (2023). Online mentoring for girls in secondary education to increase participation rates of women in STEM: A long-term follow-up study on later university major and career choices. *Annals of the New York Academy of Sciences, 1523*(1), 62–73. <https://doi.org/10.1111/nyas.14989>
- Stout, J. G., Dasgupta, N., Hunsinger, M., & McManus, M. A. (2011). Steming the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology, 100*(2), 255–270. <https://doi.org/10.1037/a0021385>
- Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics, 25*(1), 1–21. <https://doi.org/10.1214/09-STS313>
- Uebler, C., Emmendinger, K. J., Ziegler, A., & Stoeger, H. (2023). Dropping out of an online mentoring program for girls in STEM: A longitudinal study on the dynamically changing risk for premature match closure. *Journal of Community Psychology, 51*(8), 3121–3151. <https://doi.org/10.1002/jcop.23039>
- van Buuren, S., & Groothuis-Oudshoorn, K. (2006). *CRAN: Contributed packages*. <https://doi.org/10.32614/CRAN.package.mice>
- van den Hurk, A., Meelissen, M., & van Langen, A. (2019). Interventions in education to prevent STEM pipeline leakage. *International Journal of Science Education, 41*(2), 150–164. <https://doi.org/10.1080/09500693.2018.1540897>
- Vedder-Weiss, D., & Fortus, D. (2011). Adolescents' declining motivation to learn science: Inevitable or not? *Journal of Research in Science Teaching, 48*(2), 199–216. <https://doi.org/10.1002/tea.20398>
- Xplora (2004). *Xplora: European gateway to science education*. http://www.xplora.org/ww/en/pub/xplora/nucleus_home/pencil.htm
- Ziegler, A. (2009). Mentoring: Konzeptuelle Grundlagen und Wirksamkeitsanalyse [Mentoring: Conceptual foundations and effectiveness analysis]. In H. Stöger, A. Ziegler, & D. Schimke (Eds.), *Mentoring: Theoretische Hintergründe, empirische Befunde und praktische Anwendungen* (1st edition, pp. 7–30). Pabst Science Publishers.
- Ziegler, A., Forche, A. S., Graber, L., & Stoeger, H. (2025a). The capital of cooperation: A resource-based study of school–out-of-school STEM cooperations across phases and professional roles. *International Journal of STEM Education, 12*(1), 65. <https://doi.org/10.1186/s40594-025-00587-7>
- Ziegler, A., Shiani, M., Wengler, D., & Stoeger, H. (2025b). Mapping collaborations in STEM education: A scoping review and typology of in-school–out-of-school partnerships. *Education Sciences, 15*(11), 1513. <https://doi.org/10.3390/educsci15111513>

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.