How Learners Perceive Inter-Disciplinary Dialogic Science Related Learning Experiences: A Mixed-Method Study

Aviv J. Sharon[1], Lihi Telem[1], Einat Heyd-Metzuyanim[1,*]

[1] Faculty of Education in Science and Technology, Technion – Israel Institute of Technology, Haifa, Israel; einathm@technion.ac.il

Well-designed, dialogic, and interdisciplinary learning opportunities in science related disciplines can promote many educational goals. The difficulty lies in providing these opportunities in schools. Can out-of-school focus days fill the void? Here, we performed a mixed methods study to discover how middle school learners perceive focus days that included dialogic, interdisciplinary activities in mathematics, science, and philosophy, all around socio-scientific dilemmas. We also sought to understand the experiences' impact on learners' self-perceived dialogic engagement in the school classroom. We focused on (1) learners' perceptions and recollections of the dialogic climate during the focus days; (2) the perceived contribution of the three disciplines (mathematics, science and philosophy) to the interdisciplinary focus days; and (3) the focus days' effects on the dialogic climate in the everyday school environment. Our findings suggest that (1) learners generally considered the focus days as enjoyable and different than their everyday school experience; (2) students did not consider each discipline equally relevant to the focus days' main topics; and (3) the effects of the focus days were localized, that is, no impact was found on the dialogic climate at school. We discuss implications for promoting interdisciplinary, dialogic science related learning in secondary schools.

Key words: dialogic instruction, classroom dialogue, interdisciplinary approach, educational environment, student attitudes.

Received 11/2023
Revised 1/2024
Accepted 3/2024

1 Introduction

Well-designed dialogic instruction in science related disciplines can develop learners' reasoning skills and serve many other educational goals, such as promoting a deeper understanding of the scientific subjects taught at school and strengthening students' preparation for civic life in democratic societies (Clarke et al., 2016; Schwarz & Baker, 2016; Skidmore, 2019; Wegerif, 2006). However, cultivating dialogic learning in schools has encountered persistent challenges, such as those stemming from the school's institutional and regulatory environment (Hardman, 2019; Warwick & Cook, 2019).

It has also been argued that *interdisciplinarity* can promote learning and can serve as a venue for dialogism (Chettiparamb, 2007; Koichu et al., 2022). Nevertheless, like dialogic instruction, interdisciplinarity is difficult to implement in schools (Margot & Kettler, 2019). For young learners, such as middle school students, interdisciplinarity poses a unique challenge, as they are in the early stages of developing their disciplinary knowledge and identities, particularly in STEM related disciplines (Jiang et al., 2019).

Dialogic learning and interdisciplinarity in STEM education have mostly been studied from the perspectives of researchers and teachers, rather than that of learners. Very few studies have been done to understand the views of younger learners in particular. Here, we sought to understand middle school students' perceptions of an inter-disciplinary, dialogic educational intervention: out-of-school "focus days" around socio-scientific dilemmas. In the following section, we review the literature on dialogic and interdisciplinary teaching and learning, focusing on secondary education and the learners' emic perspective.

2 Theoretical Framework and Related Literature

2.1 Dialogic Learning

One of the most significant current discussions in education relates to dialogic pedagogies – a group of approaches that create opportunities for learners to express their own ideas, elaborate on them, and listen and respond to their peers (Warwick & Cook, 2019). Dialogic approaches in education strive to help learners explore unfamiliar and complex subjects and become more competent at reasoning tasks. These approaches also aim to help students recognize the value of dialogue itself, help them "become more open to learning" in general (Wegerif, 2019, p. 20) and "enlist them as active participants in the process of knowledge production" as well as in broader processes of social transformation (Skidmore, 2019, p. 35). Along similar lines, Schwarz and Baker (2016) argue that enacting dialogue "that integrates rigorous reasoning and accountability to the other" in educational settings is crucial if we are to prepare learners for citizenship in a deliberative democracy (p. 233).

Research on classroom dialogue has focused on more mainstream learning outcomes, such as conceptual learning, knowledge retention over time, and reasoning and problem-solving skills, as well as knowledge transfer to other domains. By and large, the extant body of evidence indicates that dialogic instruction can improve student learning along these dimensions (Clarke et al., 2016; Koichu et al., 2022).

2.2 Interdisciplinarity and Dialogic Learning

Interdisciplinarity requires a synthesis between bodies of knowledge and is considered crucial for solving complex problems. Its advocates have argued that it is desirable and inevitable in policy, research and education (Chettiparamb, 2007; Nowacek, 2005; Spelt et al., 2009).

Interdisciplinarity appears to be related to dialogic learning in at least two ways. First, it has been argued that fostering interdisciplinary cognition requires a dialogic context. Nikitina (2005) contended that interdisciplinary processes bring each discipline's characteristic languages, epistemologies and cultures to the foreground and make them more salient. This facilitates connections between the disciplinary ideas. She also argued, however, that building upon ideas across disciplines requires several cognitive moves. These include: (1) overcoming a monodisciplinary perspective and being receptive to other voices, while retaining one's own personal voice; (2) explicitly bridging different disciplinary perspectives to create a synthesis; and (3) systematically critiquing the provisional synthesis. Additionally, performing these moves closely corresponds with conducting a dialogue with oneself that incorporates multiple voices. Thus, the argument goes, interdisciplinary processes *require* dialogism.

Second, Koichu et al. (2022) put forward a claim in the opposite direction: namely, that dialogic learning can be fostered *by* interdisciplinarity. Underlying this claim is the premise that if learners are asked to discuss interdisciplinary problems, they are likely to find that they can resolve them only by combining ideas from different disciplines. Thus, an interdisciplinary task can *prompt* dialogism.

Prompting interdisciplinary dialogue in middle schools, particularly around science related disciplines, poses a unique challenge, since younger learners are still not well acquainted with the disciplines to begin with. Jiang et al. (2019) suggested design principles to accommodate this. Specifically, based on their study of an interdisciplinary program for STEM+ L (Science, Technology, Engineering, Math and Digital Literacies) for middle school students, the researchers found that learners must be given opportunities to develop individual expertise in specific using methods such as role-taking. Only once individual expertise is established, the argument goes, can teams of learners discuss and investigate disciplinary ideas collectively and connect them in later steps.

In conclusion, dialogism and interdisciplinarity are interrelated and can contribute to learning in many ways. Interdisciplinary learning can take place at several educational stages, but at early stages, the learners' familiarity with the disciplines constrains instructional and task design. Whether dialogism and interdisciplinarity are implemented in schools is another matter.

2.3 Challenges of Cultivating Dialogic Learning and Interdisciplinarity in Schools

Alongside the many known benefits of dialogic education, the literature suggests that "in many classrooms, and across countries, opportunities for promoting more dialogue [. . .] are often missed" (Warwick & Cook, 2019, p. 122). This appears to be the case especially in low-performing schools (Clarke et al., 2016). Some of these challenges of implementation stem from the school's institutional and regulatory environment. Specifically, Hardman (2019) has pointed out the importance of "policies, plans and institutional arrangements with regard to teacher education, curriculum reform and assessment practices" (p. 147).

Similarly, in their analysis of science teachers' instructional practices, Zafrani and Yarden (2022) identified three prevailing "institutional logics" that constrain dialogic argumentation in classroom discourse: (1) the logic of accountability: the misuse of standardized test scores by educational leadership; (2) the logic of tracking: classroom dialogue is considered most appropriate for high-track classrooms; and (3) the logic of the profession: facilitating classroom dialogue is considered inconsistent with normative teaching practices. Since facilitating dialogue goes against these logics, dialogic approaches are implemented in schools only sporadically at best.

Like dialogism, embedding interdisciplinarity in educational settings has faced significant hurdles. This research subfield is still in its early stages, with many of the studies focusing on institutions of higher education (IHEs). Chettiparamb (2007) has identified three types of problems in fostering interdisciplinarity in teaching within these institutions, some of which resemble Zafrani and Yarden's three logics: (1) "institutional problems": IHEs' organizational structures, reward systems, etc. do not lend themselves to interdisciplinary teaching, and are difficult to change; (2) "people problems": teachers, students and the public at large resist interdisciplinary teaching because they find it difficult to grasp, are unsure of its

benefits or consider its implementation too risky; (3) "facilities problems": IHEs' physical structures and timetables make it difficult to run interdisciplinary programs, for example, due to the lack of appropriate spaces.

Secondary schools face many of the same challenges in embedding interdisciplinarity as IHEs, particularly in STEM (science, technology, engineering and mathematics). According to Margot and Kettler (2019), teachers report many challenges in embedding integrative STEM approaches within schools, including inadequate administrative and financial support. For example, teachers attest that they often do not have enough planning time and opportunities to collaborate with colleagues across disciplines (which evokes Chettiparamb's "institutional problems"). Additionally, teachers have pointed out that the implementation of STEM programs could "hinder the direct instruction of science content" required by school districts (which evokes Zafrani and Yarden's "logic of accountability"; Margot & Kettler, 2019, p. 11). Likewise, educational researchers have long pointed out institutional problems of embedding interdisciplinarity in schools (Spalding, 2002). In sum, many observers seem to agree that it is difficult to embed dialogism and interdisciplinarity in schools, and for many of the same reasons.

2.4 Learners' Perceptions of Dialogic and Interdisciplinary Learning

Studies of dialogic learning have been mostly conducted from an etic point of view on learners, meaning that of an outside observer such as a researcher or a teacher. By comparison, little attention has been paid to the students' emic perspective, although the students' voice has important implications for research and practice. Nevertheless, the extant body of evidence suggests that learners endorse diverse perceptions of dialogic learning. One strand of studies has focused on the reasons learners refrain from participating in classroom talk. One of these is the belief that only those who know correct answers have the "right to speak" during a classroom discussion (Clarke, 2015). Another reason students refrain from asking spontaneous questions is the teacher's pressure to deliver content matter efficiently (Rop, 2003). By contrast, another strand of studies suggests that under favorable conditions, learners *do* engage more often in classroom dialogue, and subsequently they become more motivated in their studies and derive more joy from schooling (Chen et al., 2020; Kiemer et al., 2015). Our recent work has focused on "dialogic climate", which we define as the extent to which dialogic engagement is normative within an educational setting, such as a class or a school. Drawing on Mercer and Littleton (2007), we define dialogic engagement as participation in co-reasoning discussions, which afford speakers opportunities to share knowledge, challenge ideas, evaluate evidence and consider options "in a reasoned and equitable way" (p. 62). Preliminary findings suggest that dialogic climate can be measured to a high level of validity using a questionnaire and that in some cases, students' ratings of dialogic climate decline over the school year (Sharon et al., in process).

Similarly, very little is known about student perceptions of interdisciplinary learning. Several studies have been conducted in learners at the tertiary educational level, but here we will focus on younger learners, at the middle-school level, who are not as familiar with the disciplines. In one study, conducted in a middle school in Finland, students aged 14 through 15 took part in a problem-based interdisciplinary course related to the natural sciences, broadly defined. Then, they were asked to reflect on their experience (Eronen et al., 2019). The findings suggest that many learners found the course to be "different from the ordinary schoolwork and thus offered a much-appreciated change" (p. 273). The learners also spoke extensively about what we consider a dialogic learning experience, by referring to a "need to learn to justify their views convincingly and [...] listen to and consider the perspectives of others" (p. 271). However, they also reported not having learned much disciplinary knowledge during the course.

In another study, conducted in middle schools in China and focusing on the natural sciences, most study participants considered interdisciplinarity to be a helpful way to integrate their scientific knowledge, but some thought that engaging in it would be too challenging for them at their current educational stage. Some learners also regarded spending time on interdisciplinary problems as useless, and even risky, as these problems are not likely to appear on admissions exams (Song & Wang, 2021).

Additional studies, conducted in Israeli high schools, focused on learners who participated in interdisciplinary programs combining physics and electronics. In these studies, learners attested that the programs helped deepen their disciplinary knowledge. They also reported interest in the programs and in interdisciplinary learning. However, learners also struggled to keep up with the fast pace of instruction and the wide scope of content (Gero et al., 2022; Gero & Zach, 2014).

In conclusion, embedding interdisciplinary and dialogic learning opportunities within schools seems to be both promising and challenging at the same time. Previous work has shed light on some aspects of embedding dialogic argumentation and interdisciplinarity, based on Dialogos, an interdisciplinary, dialogic program designed for middle school learners. Specific attention has been given to the organizational and systemic aspects of the program (Koichu et al., 2022; Zafrani & Yarden, 2022). Additionally, previous work

has characterized epistemic practices that are enacted by the learners during the program (Tsemach et al., 2023). Additional research has shown that after three years of implementing Dialogos, the project herein under investigation, teachers and other stakeholders were interested in integrating dialogic education *within* the participating schools as well (Koichu et al., 2022). Yet, the *learners'* emic perspective on Dialogos remains little understood. As learners' perspectives have received little scholarly attention in these contexts, understanding them is of great theoretical and practical importance.

2.5 Research Objective and Questions

Our objective in this study was to elucidate how middle school learners perceive inter-disciplinary, dialogic focus days around socio-scientific issues. Specifically, we sought to understand:

(1) What were learners' perceptions and recollections of the dialogic climate during the focus days? Did learners perceive them as distinct from their usual learning? If so, in what ways were they considered distinct?
(2) Did learners perceive each discipline (science, mathematics and philosophy) as relevant to the focus days? If so, in what ways?
(3) What were the focus days' effects on the dialogic climate in the everyday school environment, as perceived by the learners?

3 Methods

3.1 Research Field

This study focused on a program named Dialogos,[1] designed to foster dialogic, interdisciplinary education among junior high school students. The students participating in the program attended out-of-school "focus days" about interdisciplinary societal dilemmas, such as "Should a state be allowed to mandate parents to vaccinate their children?" or "Should citizens' DNA sequence information be added to biometric passports?". These focus days took place in a science center and were designed to give students opportunities to discuss these issues in small groups, followed by whole class discussions.

The out-of-school "focus days" were designed to avoid systemic constraints that are part and parcel of working within schools. This approach was chosen as a lesson learned from the first attempt to foster dialogic argumentation within schools, which took place in Year 1 of Dialogos in "Town A" – as described in Zafrani and Yarden (2022). As their study shows, the program encountered resistance in these schools due to institutional logics, ultimately leading to the rejection of dialogic argumentation as a "legitimate means of instruction by science teachers" (p. 163). Based on the lessons learned from Town A, in the following years, the Dialogos team drafted plans for out-of-school "focus days" according to design principles described in Koichu et al. (2022). This part took place in a different research field, which we will call "Town B," over Years 2–4 of the program (2019–2022). Here, we examine learning outcomes from Town B, with the last year (2021–2022) serving as the timeframe of data collection in the current study.

In years 2–4, the Dialogos intervention was designed to give learners an opportunity to utilize knowledge in three disciplines – mathematics, science and philosophy – in service of a larger inter-disciplinary societal dilemma as described above. In this intervention, dialogism and interdisciplinarity were considered co-constructive: the dialogic climate supported interdisciplinary discourse, and interdisciplinarity created opportunities for dialogic learning. In the following section, we detail the main characteristics of the program.

Learners participated in collaborative learning groups according to their preference of two disciplines (out of the three mentioned above) along with peers from other schools. Having learners "specialize" in specific disciplines was intended to promote dialogue since all disciplinary knowledge was necessary for the resolution of the problem at hand. The program was implemented among high-track classes only, as an attempt to "work around" the logic of tracking, a lesson learned from Year 1 of the Dialogos program (see Zafrani & Yarden, 2022). Additionally, the focus days took place off the school premises, at a local science center, which solved a "facilities problem" that often arises when trying to implement interdisciplinarity in education (Chettiparamb, 2007). However, learning did occur during school hours and participation in the program was compulsory for all the students of the high-track classrooms. In that sense, the program did follow some of the logics of school, such as the need to maintain classroom discipline.

At the end of Year 2, instruction moved online using the Zoom platform due to the spread of COVID-19. Because of this change, Year 3 was designed with distance learning in mind: considering the difficulty of

[1] Additional descriptions of this program can be found in (Zafrani & Yarden, 2022; Koichu et al., 2022).

collaborating with strangers online, the work groups were composed within schools (in homeroom groups instead of mixed groups) and the focus days were mostly interdisciplinary. Thus, in Year 3, learners were divided into disciplinary groups at the beginning of the day and then were reshuffled into interdisciplinary groups to discuss the problem, while the disciplinary preparation took place in students' schools before the focus days.

With the easing of local COVID-19 restrictions, in Year 4, the project was expanded to include two more age groups – 7$^{\text{th}}$ graders and 8$^{\text{th}}$ graders. These joined the 9$^{\text{th}}$ graders who were at the time in their third year in the project. In this year, 9^{th} graders experienced only one focus day near the year's end, while each of the two other age groups experienced two focus days. The fourth year's focus days, like Year 2's focus days, took place out of school, in a local science center in Town B, but they were designed such that students worked in homeroom groups instead of mixed groups.

The activities during the focus days were designed by the research team and led by the teachers of the students from their schools. These teachers participated in a professional development program where they became familiar with the activities before the focus days, and had opportunities to give feedback and suggest changes.

3.2 Participants

The participants of the study came from four schools, which we will call Rimonim, Savyonim, Gannim, and Nofim. The study was conducted during the 2021–2022 school year (Year 4 of the program). All schools belonged to the state-secular track and were located in the same town in central Israel. These schools were characterized by a high socio-economic status, according to the Israel Ministry of Education's socio-economic index. All schools contained both a junior high school (grades 7–9, ages 12–15) and a high school (grades 10–12, ages 16–18).

The four schools participated in years 2–4 of the Dialogos program. The intervention group was composed of high-track "excellence classes", one in each age group of junior high school. The control group was composed of homeroom classes from each age group in each school that were "matched" (selected) to serve as a control for that age group's "excellence class." The matches were performed by a representative of each school's leadership team, based on similarity in overall characteristics of the students.

3.3 Research Design

We employed a mixed methods research design in accordance with the idea that mixed research allows researchers to take advantage of "complementary strengths and non-overlapping weaknesses" of different research methods (Johnson & Onwuegbuzie, 2004, p. 18). The quantitative component in this design was a questionnaire study and the qualitative one consisted of focus groups. We anticipated that the quantitative component would allow us to investigate causal questions, such as whether the intervention impacted the dialogic climate at school, and that the qualitative component would yield fine-grained data that would provide insights into students' experiences and could help us interpret the quantitative findings. We detail the design of each of these components below.

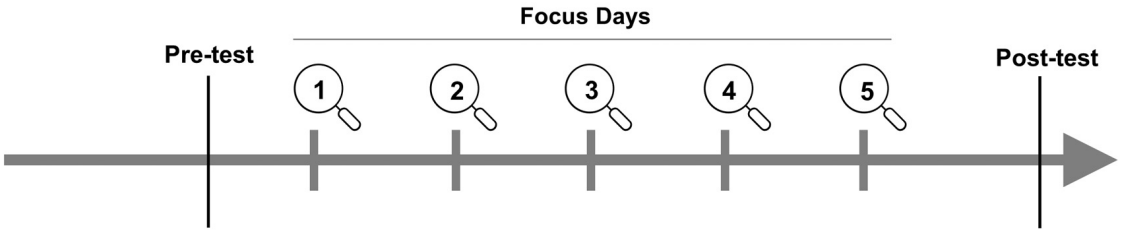
3.4 Quantitative Methods

Students enrolled in the intervention classes and the control classes were asked to fill out questionnaires in the 2021–2022 school year on several occasions: Before the intervention (pre-test; intervention and control groups; $N_{\text{intervention}} = 263$, $N_{\text{control}} = 246$), at the end of each of the five focus days (394 questionnaires in total, intervention group only), and after the intervention (post-test; intervention and control groups; $N_{\text{intervention}} = 222$, $N_{\text{control}} = 172$; Table 1; Figure 1).

Tab. 1: Sample Demographics (Pre-Test and Post-Test), by Timepoint and Condition

	Pre-Test	Post-Test	Total
Intervention	263	222	485
Control	246	172	418
Total	509	394	903

Note. Nofim questionnaires were removed from the sample because we were only able to collect post-test questionnaires at that school. By excluding these data points, we could easily compare these findings with those we derived from the paired questionnaires. n(Nofim) = 98.



Focus Day		1	2	3	4	5	
Topic	-	Ethics	Ethics	Climate	Climate	Climate	-
Grades	-	Grade 7	Grade 8	Grade 7	Grade 8	Grade 9	-
Questionnaires							
Participants	Intervent. & Control Schools R,S,G Grades 7-9	Intervent. Only Schools R,S,G,N Grade 7	Intervent. Only Schools R,S,G,N Grade 8	Intervent. Only Schools R,S,G,N Grade 7	Intervent. Only Schools R,S,G,N Grade 8	Intervent. Only Schools R,S,G,N Grade 9	Intervent. & Control Schools R,S,G,N Grades 7-9
Sample Size	510	95	83	55	95	67	394
Instrument	DCI	Focus Day Eval.	Focus Day Eval.	Focus Day Eval.	Focus Day Eval.	Focus Day Eval.	DCI
Student IDs Recorded?	Yes	No	No	No	Yes	Yes	Yes
Focus Groups							
Participants	-	-	-	-	-	Intervent. Only; Schools R,S,G,N; Grade 9	Intervent. Only; Schools R,S,G,N; Grade 9

Fig. 1: Study design. Schools R,S,G,N: Rimonim, Savyonim, Gannim, and Nofim, respectively. Intervent.: Intervention. Focus Day Eval.: Focus Day Evaluation Questionnaire

More specifically, to address RQs 1 and 2, students in the intervention group were asked to complete the Focus Day Evaluation questionnaire at the end of each focus day. These questionnaires measured and assessed (1) learners' perceptions and recollections of the dialogic climate, and (2) the perceived contribution of different disciplines. Specifically, the questionnaire focused on the following perceptions and recollections: the "Overall Experience" at the end of the focus days, on a scale of 1 ("very bad") to 5 ("really great"); a perception of "Distinctness from Usual Learning", on a scale of 1 ("very similar to what we do at school") to 5 ("the activities are very different"); "Feeling Heard" during discussions ("To what extent do you think your opinion was heard during the activities today?" on a scale of 1 to 5) and whether classroom discussions were civil and open-minded (Respectful Discourse; "In your class, to what extent did you feel there were respectful conversations that acknowledged different perspectives?" on a scale of 1 to 5). Additionally, the instrument included a "multi-disciplinary self-efficacy" index which was composed of one item each for self-efficacy in mathematics, science and language-rich subjects. (Only the intervention classes were asked to complete the focus day evaluation questionnaires, as they were the only group that took part in the focus days.)

To address RQ3, which refers to the dialogic climate within the school environment, as perceived by the students, we used another instrument, the Revised Dialogic Climate Inventory (DCI). Dialogic climate is described by three dimensions: (1) Student Voice, e.g., "When I speak in class, I explain and elaborate upon the answers and ideas that I present," I22; (2) Dialogic Atmosphere, e.g., "Students learn in this class by participating in class activities and discussions", I17; and (3) Willingness to Listen, e.g., "I can listen to others' opinions even when they go against mine", C38.

To track changes in responses over time, students were asked to enter a unique, anonymized ID number at the beginning of the questionnaire. These numbers were provided by the teachers, but not recorded by the researchers. Due to practical constraints, we were only able to pair 238 of these questionnaires between the pre-test and the post-test (Table 2). In addition, we were unable to collect pre-test questionnaires at the Nofim school, due to local COVID-19 restrictions. Finally, due to the program's constraints, we were only able to record the participants' anonymized ID numbers during Focus Days 4 and 5.

Tab. 2: Sample demographics, Paired DCI Questionnaires (Pre-Test and Post-Test), by School and Condition

Condition	Rimonim	Savyonim	Gannim	Total
Intervention	44	40	67	151
Control	26	39	22	87
Total	70	79	89	238

Much of the statistical analysis involved the use of non-parametric techniques, such as the Wilcoxon, Mann-Whitney, and Friedman tests. This selection was made since numerous metrics were derived from singular questionnaire items, and due to the advantage of these tests in making minimal assumptions about the dataset. Regarding RQ3, we employed ANOVAs with the raw data, despite occasional violations of normality. This approach was adopted since ANOVA is considered robust to violations of normality in terms of Type I error (Blanca et al., 2017), and since we had conducted a visual inspection of the relevant datasets, and we had observed no outliers and no large departures from normality.

3.5 Qualitative Methods

To address RQs 1 and 2, we conducted four focus groups among 9th graders from the intervention group, one for each of the four schools. One researcher ran all groups. Two groups were held during the last of the five focus days, and the rest were held on school grounds, during school hours, at the end of the school year. Each group was composed of four learners, and all were mixed-gender, except for one, which consisted of girls only. The focus groups participants were selected by the teachers. The discussions lasted 31 minutes (Rimonim), 54 minutes (Savyonim), 56 minutes (Ganim) and 59 minutes (Nofim). The focus group protocol and qualitative data analysis methods are described in the Appendix.

3.6 Ethics statement

Before beginning data collection, approval for the ethical aspects of this study was obtained from the Institutional Review Board of the authors' institution. Additionally, permission was obtained from the Office of the Chief Scientist of the Israel Ministry of Education.

4 Findings

We present our findings according to four themes that arose from the data. For each theme, we start by presenting the quantitative findings from the surveys, then supplement these with qualitative findings from the focus groups.

4.1 Overall Experience

Quantitative findings: Learners Generally Found the Focus Days Enjoyable

When learners were asked to rate their overall experience on a scale of 1 ("very bad") to 5 ("really great") at the end of the focus days, the overall mean score was 3.59 ($SD = 1.04$; Figure 2a). However, a different picture emerged when participants were asked how much the focus day activities differed from the usual activities the learners have at school on a scale of 1 ("very similar to what we do at school") to 5 ("the activities are very different"). Here, the mean score was slightly lower ($M = 3.38$) and there was less agreement between respondents, as evident by the higher standard deviation ($SD = 1.32$; Figure 2c). We conducted a Levene's test of equality of variances and found that the difference in variances between "Overall Experience" and "Distinctness from Usual Learning" was statistically significant: $F(1, 787) = = 29.92$, $p < 0.001$. Furthermore, two measures of effect size, $\ln VR = 0.23$ and $\ln CVR = 0.3$, suggest that there is a meaningful difference between the variability of these scores (Nakagawa et al., 2014).

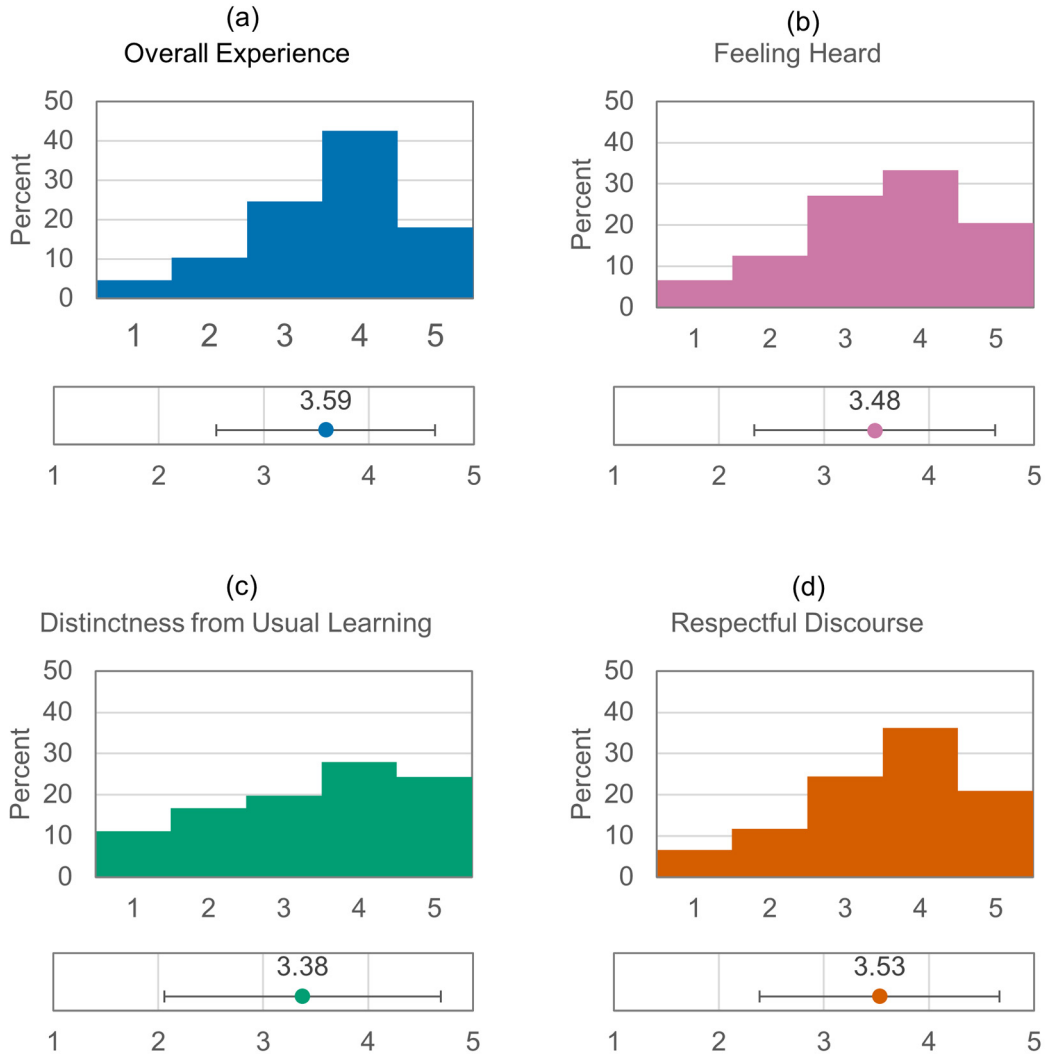


Fig. 2: Student perceptions as regards Focus Days 1-5, descriptive statistics. At the bottom of each panel, the circle represents the mean and error bars extend to one standard deviation. (a) Overall experience, $n = 394$. (b) Feeling heard, $n = 391$. (c) Distinctness from usual learning, $n = 394$. (d) Respectful discourse, $n = 392$.

Qualitative findings: Distinctness from School Learning Considered Positive

The focus groups discussions yielded a wide variety of recollections and opinions of the focus days. The participants' recollections as regards the overall experience during the focus days (RQ 1, first part) were closely related to the extent the focus days were considered distinct from learning in school (RQ 1, second and third parts). We shall thus discuss findings pertaining to the first research question in tandem.

Qualitative findings: Group Composition

A central issue that came up in all focus groups was the groups' composition during the focus days. The participants of the focus groups expressed different opinions about working with peers and teachers they did not know, as opposed to working with their own class. Some learners voiced the opinion that working with strangers made everyone more respectful and attentive, putting social status and relationships aside, so that not only popular kids had the opportunity to be heard. At the same time, working with strangers was also mentioned as inhibiting free discussion due to shyness, especially during video conferences. The following excerpt is a sample of the discussions about this topic (Table 3).

Learners also claimed that teachers had no preconceived opinions of anyone and were attentive and willing to grant permission to speak to all learners: "I think that in Dialogos we all came in as equals and not [like in class, where] the teacher remembers the same kids [who are invited to participate every time] in advance."

Tab. 3: Excerpt from Focus Group Transcript

Line	Speaker	Transcript
430	Student A	(Speaking of working with her homeroom classmates in the current focus day)... and I remember that when we were in 7th grade, with, like, other schools and kids we don't know then they had more respect for one another
431	Researcher	Aha...
432	Student A	And then suddenly, here (at the center where the focus days took place), it felt like just a lesson in class in science... suddenly the kids felt like it's just us, so they stopped you, they did...
433	Student B	They allowed themselves to interrupt
434	Student A	Yes, they... yes, I also think that they said their criticism not in the nicest way, like...
435	Researcher	Aha, when they don't know you then a more respectful discourse ensues
436	Student A	Yes, because when it was their friends, they were more respectful and they tried to be nice and not say that... and then, like, when it's me... I'm less popular, I have no problem (with that), I'm less popular in my class, and then, like, to me specifically it felt like some of the kids, did say the things, like without thinking how they are saying it.
		(Discussion continues)
449	Student A	I wasn't finished. In class it was like, it feels more comfortable to interrupt each other and that only kids that (unclear – probably "are popular") will say what they think...
450	Student C	I think that, like, when it was our class then it felt like a regular lesson and then social status has an impact...

4.2 Dialogic Climate During the Focus Days

Quantitative findings: Learners reported overall positive dialogic atmosphere during focus days

On average, over the five focus days, learners indicated that they felt that their opinion had been heard during the activities ("Feeling Heard", analogous to Student Voice: $M = 3.48$, $SD = 1.15$; Figure 2b). Learners also indicated that for the most part, they found the focus day discussions to be respectful, and that diverse opinions were considered ("Respectful Discourse", analogous to Dialogic Atmosphere: $M = 3.53$, $SD = 1.14$; Figure 2d).

We then matched questionnaires between the pre-test and focus days 4 and 5 and compared the learners' perceptions using a within-subjects design (two separate Wilcoxon tests, $n = 104$ in both cases). We found that Respectful Discourse scores, reported during the focus days, were significantly higher than the pre-test Dialogic Atmosphere scores: $Z = -2.51$, $p = 0.012 r = -0.25$. However, no statistically significant difference was found between the mean ranks of Student Voice and Feeling Heard scores: $Z = -1.43$, $p = 0.15$, $r = -0.14$ (Figure 3a).

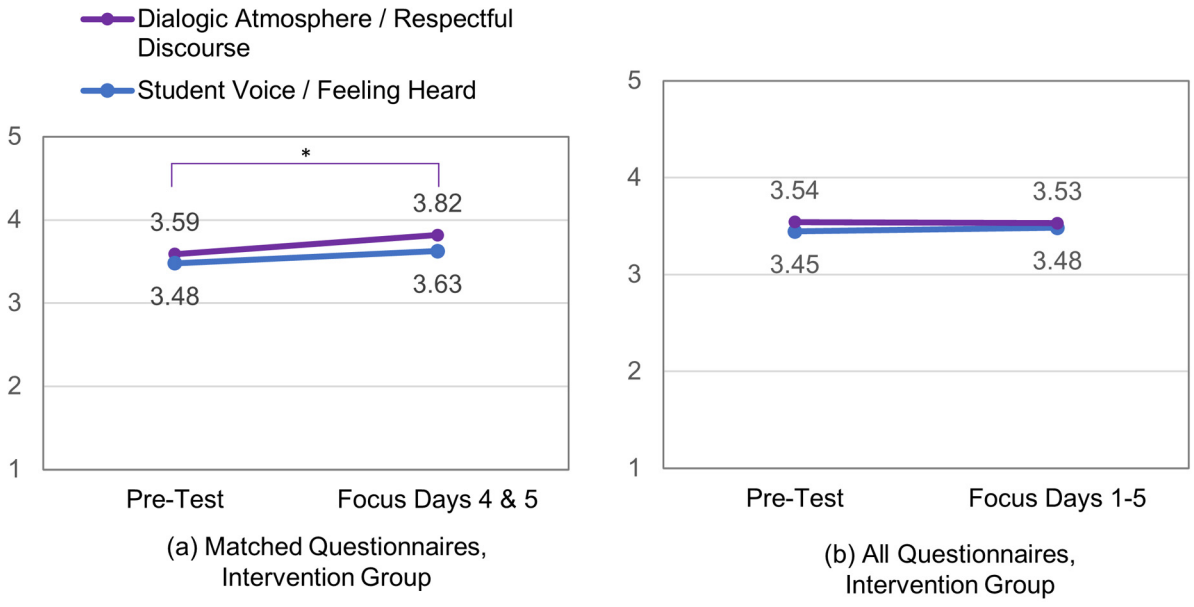


Fig. 3: (a) Dialogic climate during the pre-test vs. focus days 4 and 5 (paired data), $n = 104$. (b) Dialogic climate during the pre-test vs. focus days 1–5 (independent samples). $n(\text{pre} - \text{test}) = 485$, $n(\text{focusdays}) = 391$

We sought to validate these findings by conducting an independent-sample Mann-Whitney U tests for independent samples on a larger set of questionnaires, comparing pre-test scores with analogous scores from evaluation questionnaires from all five focus days. Using this approach, we found Dialogic Atmosphere-related scores increased over time albeit with a very small effect size, whereas the Student Voice-related scores did not (Dialogic Atmosphere/Respectful Discourse: $U = 92\,078$, $p = 0.041$, $r = = -0.07$; Student Voice/Feeling Heard: $U = 93\,359.5$, $p = 0.1$, $r = -0.05$; Figure 3b).

Quantitative findings: Self-Efficacy Correlated with Feeling Heard During Focus Days

The questionnaire responses suggest that as a learner's self-efficacy increased, so did their perception of feeling heard during the focus days. This is evident by the correlation between the multi-disciplinary self-efficacy index and "Feeling Heard" scores ($r = 0.164$, $p = 0.001$, $n = 391$). Similarly, self-efficacy in mathematics and science was also positively correlated with Feeling Heard scores: $r = 0.16$ ($p = 0.001$, $n = 391$).

Qualitative findings: Aspects of Dialogic Experience During Focus Days

In the focus group discussions, a wide variety of opinions and experiences were expressed as regards the dialogic atmosphere in the focus days. Some learners attested to a highly developed dialogic atmosphere in their school and found little difference between their regular classes and focus days. This was true specifically regarding one school, in which the students favorably pointed out Year 2 as different and more "active". Other learners reported that the dialogic climate in focus days was far more developed than in their school. However, like the dissatisfied group, they spoke positively of opportunities to be heard, of playing games and being active, as opposed to "filling out worksheets", which was perceived as schoolwork.

Focus group participants generally stated they had more opportunities to be heard and were more encouraged to listen during focus days. Some said that learning in smaller groups enabled more teacher attention and time for self-expression and some recognized the project's emphasis on dialogue. In this context, the topic of group composition came up again:

> I think that because we were with kids from different classes [schools] in seventh grade [Year 2] we felt less comfortable to interrupt people, we felt, like, a kind of commitment, sort of, to listen to them because we don't know them... but when we were with the class you felt less of such a commitment to listen to what they say, because you know them already...

When speaking of opportunities to be heard and willingness to listen in the scope of group composition, an interesting emergent theme was that of social status and its impact upon opportunities to be heard:

> When it was their friends then they were more respectful... and then when it was me... I'm less popular... to me, specifically it felt like some of the kids did say the things without thinking how they're saying it [in an offensive way].

4.3 Relevance of Disciplines

Quantitative findings: Mathematics disciplinary knowledge considered least useful for focus days

To determine whether there was any difference in the perceived relevance of the three disciplines – science, mathematics, and philosophy – to the focus day activities, we conducted a Friedman test on the questionnaire scores pertaining to this subject. According to learners' responses, mathematics was considered significantly less relevant to the focus days' main topics than science and philosophy (Science: $Mdn = 4$; Mathematics: $Mdn = 3$; Philosophy: $Mdn = 3$; $\chi^2(2) = 67.05$, $p < 0.001$, $n = 388$, Kendall's $W = 0.086$; Figure 4). The effect size ($W \leq 0.3$) indicates that there was a weak level of agreement between the learners as regards the perceived relevance of the three disciplines.

Qualitative findings: Justifications for the relevance of philosophy and science

Focus group participants used different justifications to explain why they perceived philosophy and science as more relevant and valuable to focus days projects than mathematics. Philosophy was valued for supporting and sustaining discussion, while science was considered useful for solving dilemmas. By contrast, mathematics was seen as less relevant or useful for the discussions and resolution of focus days' dilemmas. In one school, philosophy came up most strongly as more interesting and relevant than science and mathematics. It was mentioned as different from school since philosophy is not taught in schools and as being relevant and promoting discussion.

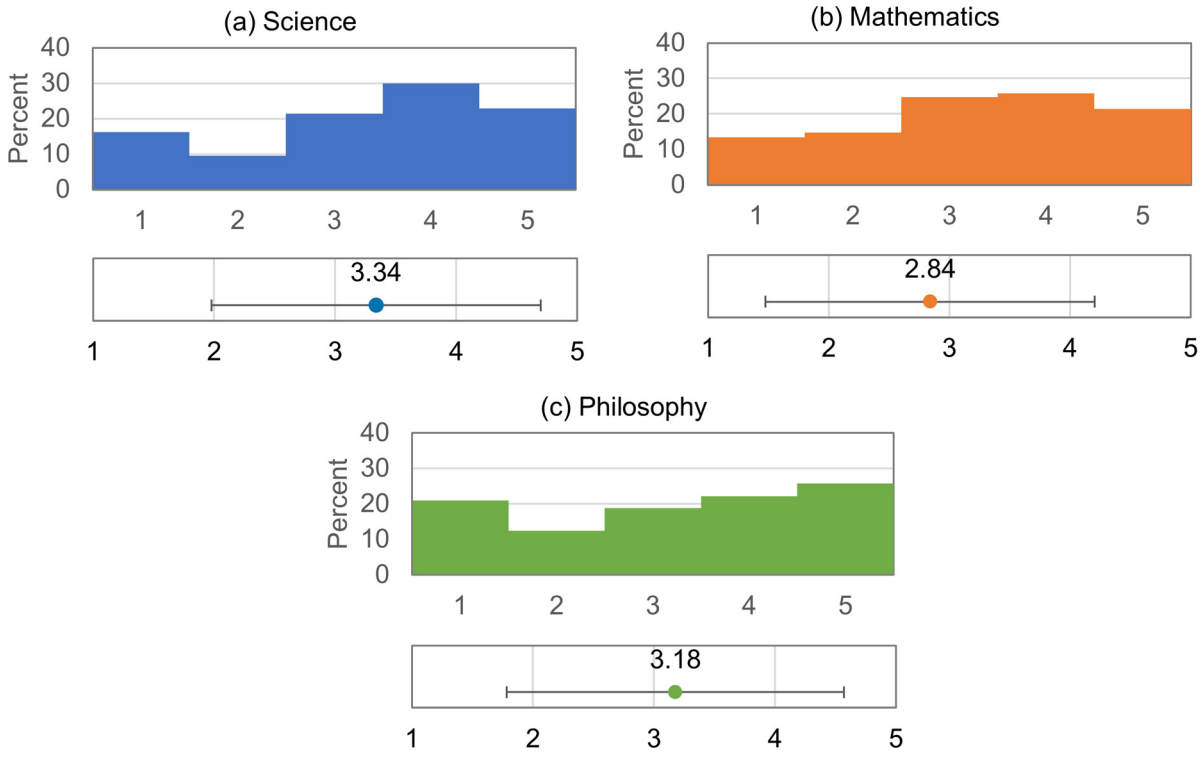


Fig. 4: Perceived usefulness of disciplinary knowledge during the focus days, by discipline, $n = 388$

4.4 Effects on Dialogic Climate in the Everyday School Environment

Quantitative findings: Dialogic climate in the school environment remained stable over time, with or without intervention

Using the DCI questionnaire data, we sought to understand the focus days' effects on the dialogic climate in the everyday school environment, as perceived by the learners. When comparing the scores from paired questionnaires (with identical student IDs in the pre-test and the post-test), the responses suggest that the dialogic climate remained stable over the school year, regardless of the intervention (Table 4 and Figure 5a). When we conducted this analysis within each age group separately, we obtained null results as well (Time × Treatment × Age Group Interaction: Student Voice: $F(2, 232) = 1.33$, $p = 0.27$; Dialogic Atmosphere: $F(2, 232) = 0.10$, $p = 0.907$; Willingness to Listen: $F(2, 232) = 0.62$, $p = 0.54$; Figure S1). We sought to validate these findings by analyzing a larger set of questionnaires, by using all valid questionnaires from the pre-test and the post-test and treating them as two independent samples. Null results were obtained when conducting this analysis as well (Table 5 and Figure S2).

Tab. 4: Associations between Condition and Dialogic Climate over Time – Two-Way Mixed ANOVA findings (paired DCI data, $\alpha = 0.05$)

Dependent Variable	Is the variance in the dependent variable significantly associated with…?		
	Time (within-subjects; Pre-test vs. Post-test)	Condition (between-subjects; Intervention vs. Control)	Time × Condition (Interaction Effect)
Student Voice	No $F(1, 236) = 1.024$, $p = 0.31$	Yes $F(1, 236) = 8.72$, $p = 0.003$ partial $\eta_2 = 0.036$	No $F(1, 236) = 0.28$, $p = 0.60$
Dialogic Atmosphere	No $F(1, 236) = 0.88$, $p = 0.35$	No $F(1, 236) = 2.15$ $p = 0.14$	No $F(1, 236) = 0.67$ $p = 0.41$
Willingness to Listen	No $F(1, 236) = 0.36$, $p = 0.55$	No $F(1, 236) = 3.10$, $p = 0.08$	No $F(1, 236) = 0.23$, $p = 0.63$

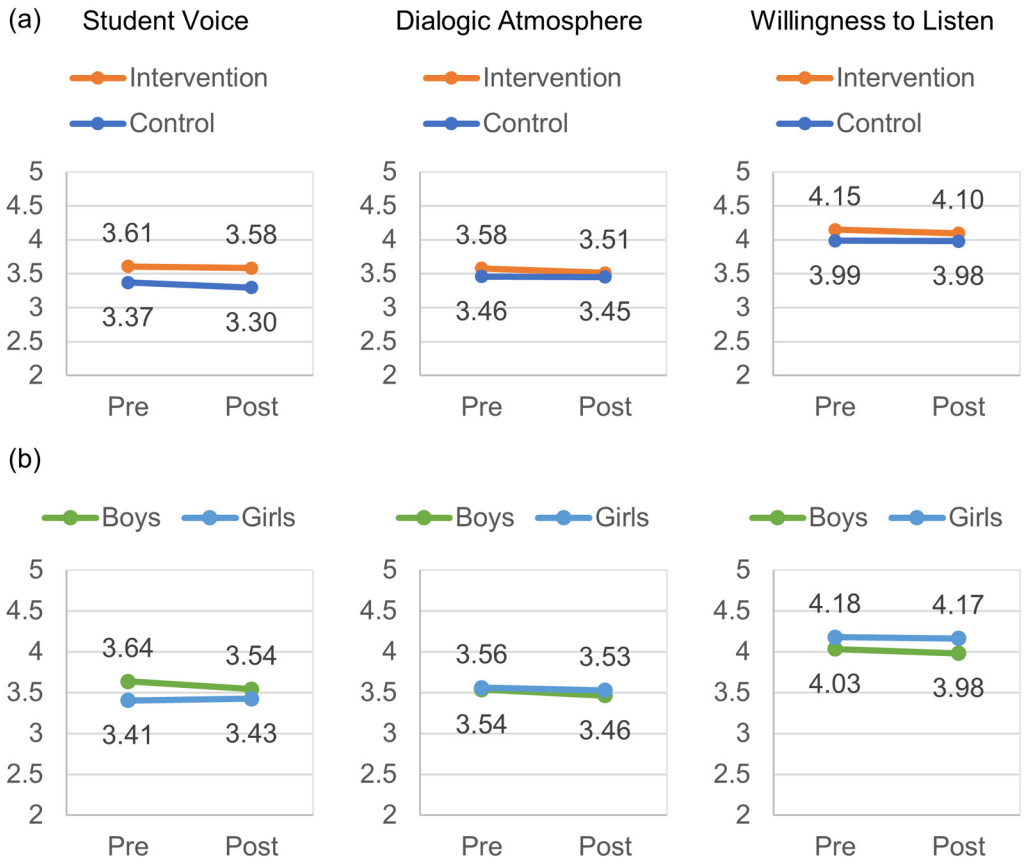


Fig. 5: Dialogic climate in the homeroom classes over time, paired data. (a) By condition (intervention vs. control). (b) By gender

Tab. 5: Associations between Condition and Dialogic Climate over Time – Two-Way ANOVA findings (All DCI Questionnaires Except Nofim, $\alpha = 0.05$)

Dependent Variable	Is the variance in the dependent variable significantly associated with...?		
	Time (between-subjects; Pre-test vs. Post-test)	Condition (between-subjects; Intervention vs. Control)	Time × Condition (Interaction Effect)
Student Voice	No $F(1, 899) = 0.01$, $p = 0.91$	Yes $F(1, 899) = 9.07$, $p = 0.003$ partial $\eta_2 = 0.01$	No $F(1, 899) = 0.78$, $p = 0.38$
Dialogic Atmosphere	No $F(1, 899) = 2.69$, $p = 0.10$	Yes $F(1, 899) = 5.72$, $p = 0.017$ partial $\eta_2 = 0.01$	No $F(1, 899) = 1.34$, $p = 0.25$
Willingness to Listen	No $F(1, 899) = 3.07$, $p = 0.08$	Yes $F(1, 899) = 8.04$, $p = 0.005$ partial $\eta_2 = 0.01$	No $F(1, 899) = 0.63$, $p = 0.80$

Quantitative findings: Learner reports of dialogic climate in the school environment varied by gender

The matched DCI questionnaire responses also reveal some divergent patterns along lines of gender in the school environment, that remained consistent throughout the year. Namely, boys reported slightly higher Student Voice scores, indicating they tended to feel that they were heard in classroom discussions, and girls reported slightly higher Willingness to Listen scores (see Table S6, Figure 5b). Similar findings were obtained when conducting this analysis using all valid questionnaires from the pre-test and the post-test and analyzing them as independent samples (Table S10, Figure S3).

The relationship between gender and dialogic climate was not raised in the focus groups. Hence, there were no qualitative findings available that could help to explain this finding.

5 Discussion

In this study, we used a mixed-method approach to understand dialogism and interdisciplinarity in discussions of socio-scientific dilemmas from the learners' emic perspective. We focused on understanding the experiences of middle school students who took part in *Dialogos*, an inter-disciplinary program designed to foster dialogic learning. Out-of-school "focus days" constituted the program's main component, and were designed to circumvent at least some of the "institutional logics" that constrain the adoption of dialogic pedagogies in schools (Zafrani & Yarden, 2022). The specific research questions pertained to: (1) learners' perceptions and recollections of the dialogic climate during the focus days; (2) learners' perceptions of the disciplines' relevance to the focus days; and (3) the focus days' effects on the dialogic climate in the school environment, as perceived by the learners.

Our main findings are as follows: (1) Learners reported an overall positive experience and found focus days to be distinct from usual schooling. Qualitative findings suggest that students considered a dialogic learning experience to be enjoyable and distinct from usual schoolwork if it required them to be active, and if it gave them an opportunity to interact with peers from other schools. (2) Learners considered science and philosophy to be more relevant than mathematics in this interdisciplinary intervention. Focus group discussions suggest that these two disciplines were perceived to be directly useful for supporting and sustaining discussion and for solving dilemmas. (3) The focus days' effect on dialogic climate did not spill over to the school environment itself, according to the students. Instead, learner reports of dialogic climate in the school environment remained stable over time, in both intervention and control groups. Additionally, responses varied slightly depending on the respondents' gender.

In the remainder of this section, we discuss the findings in light of the scholarly literature. We also call attention to limitations that should be considered when interpreting the findings. Finally, we point out the main research contributions and suggest directions for future studies.

5.1 RQ1: Learners' Perceptions and Recollections of the Dialogic Climate During the Focus Days

On average, students reported having more opportunities to speak during the focus days, compared to during their usual schoolwork. A sense of distinctness from the usual schoolwork was also reported in Eronen et al.'s study (2019) of learners who designed and built an interdisciplinary escape room at school. Similarly, previous studies have shown that when learners experience more dialogic forms of learning, they tend to report greater enjoyment in class and increased intrinsic motivation (Chen et al., 2020; Kiemer et al., 2015). This outcome is worth celebrating, given that in "business-as-usual" learning, students' interest, motivation and attitudes toward school tend to decline over time, especially as regards STEM subjects (Pekrun et al., 2017; Potvin & Hasni, 2014).

Alarmingly, our findings also suggest that as learners reported higher self-efficacy, they also tended to report having more opportunities for dialogue. Arguably, this could mean that the intervention perpetuates or even widens pre-existing inequities among learners. Since this finding is entirely based on self-reports, we do not know if this was indeed the case here. Additionally, it is not clear whether it is an effect of the intervention. Nevertheless, this finding should raise a red flag. It is possible that even within high-achieving classes, certain learners have endorsed the idea that they have less of a "Right to Speak" than their peers (Clarke, 2015) and therefore miss out on opportunities to learn. It is also possible that teachers reinforce this notion, perhaps guided by institutional logics that constrain dialogue (Zafrani & Yarden, 2022). Future work could pursue this avenue of research and suggest ways to address inequities in dialogic education.

5.2 RQ2: Learners' Perceptions of the Disciplines' Relevance to the Focus Days

Learners considered mathematics less relevant than science and philosophy in the context of the focus days. We propose two explanations for this finding. The first derives from Nikitina's (2005) argument that each discipline can be considered broadly as a culture, epistemology and language. Arguably, the epistemology and language of mathematics are relatively dissimilar to those called for when reasoning about interdisciplinary socio-scientific dilemmas. This argument is supported by Koichu et al.'s (2022) report that it was difficult to include mathematics during the design of instruction and tasks for the Dialogos focus days, compared with philosophy and science.

The second explanation for these findings derives from Jiang et al.'s (2019) notion that disciplinary knowledge is poorly developed among young learners. Specific strategies such as role-taking must be employed so these learners make the most of interdisciplinary learning activities. Although the Dialogos focus days did include some role-taking, it was not the main or only pedagogical tool used, and this

may have impacted learners' relevance assessments. This finding raises the need to more deeply consider how mathematics can be made relevant for students in real-life situations. Current developments in using mathematical tasks that encourage such connections between mathematics and real-world technologies (e.g. autonomic cars or satellite communications) can provide such connections (Kohen & Orenstein, 2021).

5.3 RQ3: Focus Days' Effects on the Dialogic Climate in the School Environment

The intervention's effects on everyday dialogic climate within schools were not statistically significant, according to student self-reports. However, there was also no statistically significant decline in dialogic climate scores, unlike past work that has shown a case in which dialogic climate scores declined over the school year (Sharon et al., in process). In any case, two explanations might help us make sense of these findings.

First, using the theoretical frameworks of Chettiparamb (2007) and Zafrani and Yarden (2022), we can speculate why Dialogos focus days were not able to foster a more dialogic climate in schools. To its credit, Dialogos seems to have circumvented some of the institutional and facilities-related problems that stand in the way of fostering dialogic and interdisciplinary learning. Arguably, however, the dialogic climate did not change because of lingering "institutional problems" and "people problems". For example, although local stakeholders expressed interest in integrating dialogism within the schools, there was only so much they could do to support and compensate teachers for their added effort. Additionally, while the solution to the "facilities problem" created new opportunities, it also meant that the school, as an institution, was less involved. Nevertheless, any attempt to generalize these findings should be done with close attention to context.

Second, these null results could also be explained by the limitations of the instrument. It may take a drastic change in classroom interactions for students to notice that the dialogic climate has changed. Hence, the DCI might not be the appropriate instrument for measuring minute effects of an out-of-school intervention. Future work could use additional methods, such as classroom observations, to address the limitations of questionnaires.

Gender and Dialogic Climate in the School Environment

Finally, we find it important to note the gender differences in the perceived dialogic climate in the classroom: boys' scores relating to their opportunities to speak were higher than girls' scores. Additionally, girls were more likely to indicate that they were willing to listen to their peers' contributions. Since these are self-reports, they do not necessarily reflect the discourse in class. However, these findings align with established knowledge about STEM-related gender stereotypes and their effects on young learners' attitudes and aspirations (Archer et al., 2014; Goetz et al., 2013). It is interesting to note that differences recurred in this study along gender lines, although the Dialogos program also included one of the humanities (philosophy). Nevertheless, this finding should raise a red flag, as it could mean that certain dialogic interventions reproduce or widen pre-existing inequities. In our view, integrating dialogism with equity and inclusion constitutes an important direction for future research.

6 Limitations and Contribution

On top of the limitations noted above, it is worth noting a few more limitations that should be considered when interpreting the findings. The first set of limitations relates to representativeness. Due to practical constraints, the intervention was conducted in ability-tracked "excellence classes", which limits what one can generalize from the study. Additionally, within the intervention group, focus groups were conducted only among 9^{th} graders, because they were most familiar with the Dialogos program. All focus group participants were also chosen by the teachers. Taken together, although these students' statements were highly informative, these sampling limitations could mean that certain learners' perspectives were not represented.

A second set of limitations relates to the study design used to address RQ3. We chose a quasi-experimental, non-equivalent group design, again due to practical constraints. Therefore, the findings may be affected by confounding variables that were not accounted for in the study.

Third, while the students' perceptions of dialogism and interdisciplinarity are valuable for theoretical and practical purposes, naturally, they do not tell the whole story. Future studies of dialogic climate should solicit input from additional stakeholders, such as teachers, principals, and parents.

Despite these limitations, our findings shed light on the affordances and limitations of integrating dialogism and interdisciplinarity within secondary schools. Dialogue and interdisciplinarity are central

to the development of competent and engaged citizens in the 21st century. However, multiple studies have shown the challenges of incorporating these elements into "mainstream" school classrooms. This study deepens the understanding of the affordances and limitations of integrating dialogism and interdisciplinarity within out-of-school programs. It illustrates the challenges of designing a program that could affect how students perceive their own opportunities to participate and to be heard. In our opinion, such findings should not discourage educators from designing such interventions, but rather nudge them towards being more realistic about their expected outcomes.

Acknowledgment

The study was supported by the Israel Science Foundation, grant no. 2699/17.

References

Archer, L., DeWitt, J., & Willis, B. (2014). Adolescent boys' science aspirations: Masculinity, capital, and power. *Journal of Research in Science Teaching*, *51*(1), 1–30. https://doi.org/10.1002/tea.21122

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, *29*(4), 552–557. https://doi.org/10.7334/psicothema2016.383

Chen, G., Zhang, J., Chan, C. K. K., Michaels, S., Resnick, L. B., & Huang, X. (2020). The link between student-perceived teacher talk and student enjoyment, anxiety and discursive engagement in the classroom. *British Educational Research Journal*, *46*(3), 631–652. https://doi.org/10.1002/berj.3600

Chettiparamb, A. (2007). *Interdisciplinarity: A literature review*. The Interdisciplinary Teaching and Learning Group, Subject Centre for Languages, Linguistics and Area Studies, School of Humanities, University of Southampton.

Clarke, S. N. (2015). The right to speak. In *Socializing intelligence zhrough academic talk and dialogue* (pp. 167–180). American Educational Research Association.

Clarke, S. N., Resnick, L. B., & Penstein Rosé, C. (2016). Dialogic instruction: A new frontier. In L. Corno, & E. M. Anderman (Eds.), *Handbook of educational psychology* (3rd ed., pp. 378–389). Routledge.

Eronen, L., Kokko, S., & Sormunen, K. (2019). Escaping the subject-based class: A Finnish case study of developing transversal competencies in a transdisciplinary course. *The Curriculum Journal*, *30*(3), 264–278. https://doi.org/10.1080/09585176.2019.1568271

Gero, A., Essami, H., Danino, O., & Kornblum, L. (2022). Students' attitudes toward interdisciplinary learning: A high-school course on solar cells. *International Journal of Engineering Education*, *38*(4), 1130–1140.

Gero, A., & Zach, E. (2014). High school programme in electro-optics: A case study on interdisciplinary learning and systems thinking. *International Journal of Engineering Education*, *30*(5), 1190–1199.

Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R., & Hall, N. C. (2013). Do girls really experience more anxiety in mathematics? *Psychological Science*, *24*(10), 2079–2087. https://doi.org/10.1177/0956797613486989

Hardman, F. (2019). Embedding a dialogic pedagogy in the classroom: What is the research telling us? In *The Routledge international handbook of research on dialogic education* (pp. 139–151). Routledge. https://doi.org/10.4324/9780429441677-13

Jiang, S., Shen, J., & Smith, B. E. (2019). Designing discipline-specific roles for interdisciplinary learning: two comparative cases in an afterschool STEM+L programme. *International Journal of Science Education*, *41*(6), 803–826. https://doi.org/10.1080/09500693.2019.1581958

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14–26. https://doi.org/10.3102/0013189X033007014

Kiemer, K., Gröschner, A., Pehmer, A. K., & Seidel, T. (2015). Effects of a classroom discourse intervention on teachers' practice and students' motivation to learn mathematics and science. *Learning and Instruction*, *35*, 94–103. https://doi.org/10.1016/j.learninstruc.2014.10.003

Kohen, Z., & Orenstein, D. (2021). Mathematical modeling of tech-related real-world problems for secondary school-level mathematics. *Educational Studies in Mathematics*, *107*, 71–91. https://doi.org/10.1007/s10649-020-10020-1

Koichu, B., Schwarz, B. B., Heyd-Metzuyanim, E., Tabach, M., & Yarden, A. (2022). Design practices and principles for promoting dialogic argumentation via interdisciplinarity. *Learning, Culture and Social Interaction*, *37*(July 2021), 100657. https://doi.org/10.1016/j.lcsi.2022.100657

Margot, K. C., & Kettler, T. (2019). Teachers' perception of STEM integration and education: a systematic literature review. *International Journal of STEM Education*, *6*(2). https://doi.org/10.1186/s40594-018-0151-2

Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. Routledge.

Nakagawa, S., Poulin, R., Mengersen, K., Reinhold, K., Engqvist, L., Lagisz, M., & Senior, A. M. (2014). Meta-analysis of variation: Ecological and evolutionary applications and beyond. *Methods in Ecology and Evolution*, *6*(2), 143–152. https://doi.org/10.1111/2041-210x.12309

Nikitina, S. (2005). Pathways of interdisciplinary cognition. *Cognition and Instruction*, *23*(3), 389–425. https://doi.org/10.1207/s1532690xci2303_3

Nowacek, R. S. (2005). A discourse-based theory of interdisciplinary connections. *The Journal of General Education*, *54*(3), 171–195. https://doi.org/10.1353/jge.2006.0006

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development*, *88*(5), 1653–1670. https://doi.org/10.1111/cdev.12704

Potvin, P., & Hasni, A. (2014). Interest, motivation and attitude towards science and technology at K-12 levels: a systematic review of 12 years of educational research. *Studies in Science Education*, *50*(1), 85–129. https://doi.org/10.1080/03057267.2014.881626

Rop, C. J. (2003). Spontaneous inquiry questions in high school chemistry classrooms: Perceptions of a group of motivated learners. *International Journal of Science Education*, *25*(1), 13–33. https://doi.org/10.1080/09500690210126496

Schwarz, B. B., & Baker, M. J. (2016). *Dialogue, argumentation and education: History, theory, and practice*. Cambridge University Press.

Sharon, A. J., Ben-Dor, N., & Heyd-Metzuyanim, E. (in process). *Conceptualizing and measuring dialogic climate: The Dialogic Climate Inventory (DCI)* [Manuscript in preparation]. Faculty of Education in Science and Technology, Technion – Israel Institute of Technology.

Skidmore, D. (2019). Dialogism and education. In N. Mercer, R. Wegerif, & L. Major (Eds.), *The Routledge international handbook of research on dialogic education* (pp. 27–37). Routledge.

Song, G., & Wang, Z. (2021). Factors influencing middle school students' interdisciplinary competence in science education. *Journal of Research in Science Teaching*, *58*, 1041–1072. https://doi.org/10.1002/tea.21692

Spalding, E. (2002). Of organelles and octagons: What do preservice secondary teachers learn from interdisciplinary teaching? *Teaching and Teacher Education*, *18*(6), 699–714. https://doi.org/10.1016/S0742-051X(02)00029-X

Spelt, E. J. H., Biemans, H. J. A., Tobi, H., Luning, P. A., & Mulder, M. (2009). Teaching and learning in interdisciplinary higher education: A systematic review. *Educational Psychology Review*, *21*(4), 365–378. https://doi.org/10.1007/s10648-009-9113-z

Tsemach, E., Schwarz, B. B., Israeli, M., & Keynan, O. (2023). Advancing group epistemic practices in the resolution of interdisciplinary societal dilemmas. *Dialogic Pedagogy*, *11*(3), A119–A147. https://doi.org/10.5195/dpj.2023.551

Warwick, P., & Cook, V. (2019). Classroom dialogue. In N. Mercer, R. Wegerif, & L. Major (Eds.), *The Routledge international handbook of research on dialogic education* (pp. 121–124). Routledge. https://doi.org/10.4324/9780429441677-11

Wegerif, R. (2006). Dialogic education: What is it and why do we need it? *Education Review*, *19*(2), 58–66.

Wegerif, R. (2019). Towards a dialogic theory of education for the internet age. In N. Mercer, R. Wegerif, & L. Major (Eds.), *The Routledge international handbook of research on dialogic education* (pp. 14–26). Routledge.

Zafrani, E., & Yarden, A. (2022). Dialog-constraining institutional logics and their interactional manifestation in the science classroom. *Science Education*, *106*(1), 142–171. https://doi.org/10.1002/sce.21687

Appendix

Quantitative Data Collection

The Focus Day Evaluation Questionnaire

The evaluation questionnaire contained Likert-scale and open-response items designed to assess perceptions and recollections of the dialogic climate (RQ1) and the perceived contribution of different disciplines to the learning experience (RQ2). The questionnaire included one five-point Likert-scale item for each

of the following constructs: (1) the learners' overall experience during the activities; (2) the extent they felt heard (similar to the Student Voice scale in the DCI, see below); (3) the extent they perceived the discourse as respectful (similar to the Dialogic Atmosphere scale in the DCI); and (4) distinctness from usual learning ("To what extent did you feel that the focus day activities were different from what you usually do at school?").

Additional items related to learners' perceptions of: (5) the best activity during the focus day; (6) the preparation activities; (7) the learners' own self-efficacy in mathematics, science and language-rich subjects (such as literature, history, and bible studies); (8) the usefulness of the knowledge they had gained at school in each of the disciplines with respect to the focus days; and (9) the Dialogos project in general.

Finally, the instrument included a "multi-disciplinary self-efficacy" index which was composed of one item each for self-efficacy in mathematics, science and language-rich subjects. This index had a Cronbach's α value of 0.73.

The Revised Dialogic Climate Inventory (DCI)

The Dialogic Climate Inventory (Sharon et al., in preparation) is designed to measure the dialogic climate in the everyday school environment, as perceived by the learners (Table S1). The original version of the DCI contained 19 Likert scale items, divided into three factors: Student Voice, Dialogic Atmosphere and Willingness to Listen, containing 7, 9 and 3 items, respectively (Sharon et al., in preparation). Some example items are: "When I speak in class, I explain and elaborate upon the answers and ideas that I present" (I22, Student Voice); "Students learn in this class by participating in class activities and discussions" (I17, Dialogic Atmosphere) and "I can listen to others' opinions even when they go against mine" (C38, Willingness to Listen).

Here, we revised the DCI by adding an extra item to the Willingness to Listen factor to improve the internal consistency of this construct. The added item was: "Even when I hear an opinion that differs from my own, I listen to it patiently until the end" (C42). Thus, the instrument we used contained 20 items in total, which were presented to the participants in random order. In this study, the Cronbach's Alpha values observed for each factor were: Student Voice (7 items): .86; Dialogic Atmosphere (9 items): 0.77; Willingness to Listen (4 items): 0.69 (pre-test data, $n = 510$).

Quantitative Data Analysis

To investigate RQ1, we computed descriptive statistics for several items of the focus day evaluation questionnaires. When possible, we also sought to discern changes between the pre-test and the focus days by matching between questionnaires based on anonymized ID numbers and employing several statistical tests. Specifically, we compared the dialogic climate during the pre-test with that of focus days 4 and 5. We retrieved Student Voice and Dialogic Atmosphere scores from the revised DCI and compared them with the Feeling Heard and Respectful Discourse scores from the focus day evaluation questionnaires, respectively. The statistical analysis was then performed using dependent-sample t-tests.

Next, we sought to validate this method since the paired sample was small (104 questionnaires). Hence, we repeated this analysis by using all valid questionnaires and conducting independent-sample t-tests.

To address RQ2, we computed descriptive statistics for the scores indicating the perceived contribution of different disciplines to the focus days. We then conducted a Friedman's test to examine whether there were any differences in the responses.

To address RQ3, we sought to discern whether the focus days affected the dialogic climate in the everyday school environment. To that end, we matched between pre- and post-test questionnaires using the anonymized ID numbers and conducted mixed two-way ANOVAs, using the DCI dimensions as dependent variables. The independent variables were condition (intervention vs. control; between-subjects) and time (pre-test vs. post-test; within-subjects).

Additionally, we sought to validate this method since the paired sample was small (238 questionnaires) and the control group was under-represented (87 questionnaires, compared to 151 in the intervention group). Hence, we repeated this analysis by using all valid DCI questionnaires and conducting between-subjects ANOVAs.

Similarly, to investigate whether gender explained variance in dialogic climate over time, we ran both mixed and between-subjects ANOVAs, using gender and time as independent variables.

Focus Group Protocol

The researcher introduced herself, explained the purpose of the group as part of the research study and informed the participants that they were free to stop participating at any time. Then, the researcher proceeded to collect the participants' names, schools, and ages.

The discussion was semi-structured and focused on the following questions:

1. What was your general impression and experience in the project?
2. What was the best focus day, and why?
3. Were there *Dialogos* activities in school? Which ones?
4. Was the learning during the focus days different from learning at school?
5. Could you compare face-to-face focus days with focus days over Zoom?
6. Which discipline was most significant for the focus days?
7. Were there *Dialogos*-style lessons in school? How were they different from regular lessons?
8. Did you feel you had the opportunity to be heard in the focus day? Thanks to what?
9. What did you learn in the *Dialogos* project?
10. Would you recommend the program? To which kind of learners?

All group discussions were recorded and transcribed. For the analysis, transcriptions were coded both top-down (for specific references to the topics of the research questions) and bottom-up (for emergent themes). Then the codes were clustered, and their content was analyzed thematically.

Tab. S1: Items of the revised Dialogic Climate Inventory (DCI)

	Item	Label
		Student Voice
1.	I20	In classroom discussions, I talk and share my ideas.
2.	I18	I like participating in discussions very much.
3.	I22	When I talk in class, I explain and elaborate on my answers and ideas.
4.	I30	I feel that I can present new ideas in class.
5.	I26	I feel comfortable participating and speaking in class.
6.	C35	I feel that my opinion has an impact in class.
7.	C33	I feel that I can persuade others even when my opinion differs from theirs.
		Dialogic Atmosphere
8.	I23	The students in my class actively participate in lessons and tasks.
9.	I27	Each student in class has the opportunity to share their ideas and thoughts.
10.	I28	When a student presents an answer or an idea, the other students listen.
11.	I29	The students in class present different opinions during tasks and activities.
12.	I17	Students learn in my class by participating in class activities and discussions.
13.	I31	In my class, students are responsible for their own learning and progress.
14.	I19	Class discussions help me think better.
15.	I21	In class, students feel comfortable admitting and correcting the mistakes they made during their learning.
16.	C40	Having different opinions in class is good because you can learn from everyone.
		Willingness to Listen
17.	C39	I get angry and shout when children in class express opinions that go against mine (reverse coded).
18.	C38	I can listen to others' opinions even when they go against mine.
19.	C41	There are opinions in class that I'm not willing to listen to at all (reverse coded).
20.	C42	Even when I hear an opinion that differs from my own, I listen to it patiently until the end.

Tab. S2: Sample demographics, Focus Days 1–5, by School and Age Group

	Focus Days	Rimonim	Savyonim	Gannim	Nofim	Total
Grade 7	1, 3	30	44	34	41	149
Grade 8	2, 4	42	43	50	43	178
Grade 9	5	24	11	20	12	67
Total	–	96	98	104	96	394

Note: Sample points refer to the total numbers of questionnaires, not unique participants.

Tab. S3: Sample demographics, Paired Questionnaires, Pre-Test DCI and Focus Days 4 & 5, by School and Age Group

	Focus Day	Rimonim	Savyonim	Gannim	Total
Grade 8	4	18	27	13	58
Grade 9	5	17	9	20	46
Total	–	35	36	33	104

Tab. S4: Sample demographics, Paired DCI Questionnaires (Pre-Test and Post-Test), by Condition and Gender

	Boys	Girls	Other / Prefer not to answer	Total
Intervention	92	56	3	151
Control	41	44	2	87
Total	133	100	5	238

Tab. S5: Sample demographics, Paired DCI Questionnaires (Pre-Test and Post-Test), by Condition and Age Group

	Grade 7	Grade 8	Grade 9	Total
Intervention	58	58	35	151
Control	21	32	34	87
Total	79	90	69	238

Tab. S6: Associations between time, gender and dialogic climate – Two-Way Mixed ANOVA findings (paired DCI data, $\alpha = 0.05$)

Dependent Variable	Is the variance in the dependent variable significantly associated with...?		
	Time (within-subjects; Pre-test vs. Post-test)	Gender (between-subjects; Boys vs. Girls)	Time × Gender (Interaction Effect)
Student Voice	No $F(1, 231) = 0.60$, $p = 0.44$	Yes $F(1, 231) = 4.10$, $p = 0.04$	No $F(1, 231) = 1.33$, $p = 0.25$
Dialogic Atmosphere	No $F(1, 231) = 1.85$, $p = 0.18$	No $F(1, 231) = 0.58$, $p = 0.45$	No $F(1, 231) = 0.30$, $p = 0.58$
Willingness to Listen	No $F(1, 231) = 0.48$, $p = 0.49$	Yes $F(1, 231) = 4.48$, $p = 0.04$	No $F(1, 231) = 0.14$, $p = 0.71$

Tab. S7: Sample demographics (Pre-Test and Post-Test), by Condition and Age Group, all DCI Questionnaires except Nofim

	Grade 7	Grade 8	Grade 9	Total
Intervention	182	169	134	485
Control	135	152	131	418
Total	317	321	265	903

Tab. S8: Sample Demographics (Pre-Test and Post-Test), by School and Condition, all DCI Questionnaires Except Nofim

	Rimonim	Savyonim	Gannim	Total
Intervention	156	154	175	485
Control	126	169	123	418
Total	282	323	298	903

Tab. S9: Sample demographics (Pre-Test and Post-Test), by Condition and Gender, all DCI Questionnaires except Nofim

	Boys	Girls	Other / Prefer not to answer	Total
Intervention	279	186	20	485
Control	198	209	11	418
Total	477	395	31	903

Tab. S10: Associations between time, gender and dialogic climate – Two-Way ANOVA findings (all DCI questionnaires except Nofim, $\alpha = 0.05$)

Dependent Variable	Is the variance in the dependent variable significantly associated with...?		
	Time (between-subjects; Pre-test vs. Post-test)	Gender (between-subjects; Boys vs. Girls)	Time × Gender (Interaction Effect)
Student Voice	No $F(1, 869) = 0.05$, $p = 0.83$	Yes $F(1, 869) = 5.55$, $p = 0.02$	No $F(1, 869) = 1.09$, $p = 0.30$
Dialogic Atmosphere	No $F(1, 869) = 2.86$, $p = 0.09$	Yes $F(1, 869) = 9.95$, $p = 0.002$	No $F(1, 869) = 0.32$, $p = 0.57$
Willingness to Listen	No $F(1, 869) = 3.25$, $p = 0.07$	Yes $F(1, 869) = 22.32$, $p < 0.001$	No $F(1, 869) = 0.83$, $p = 0.36$

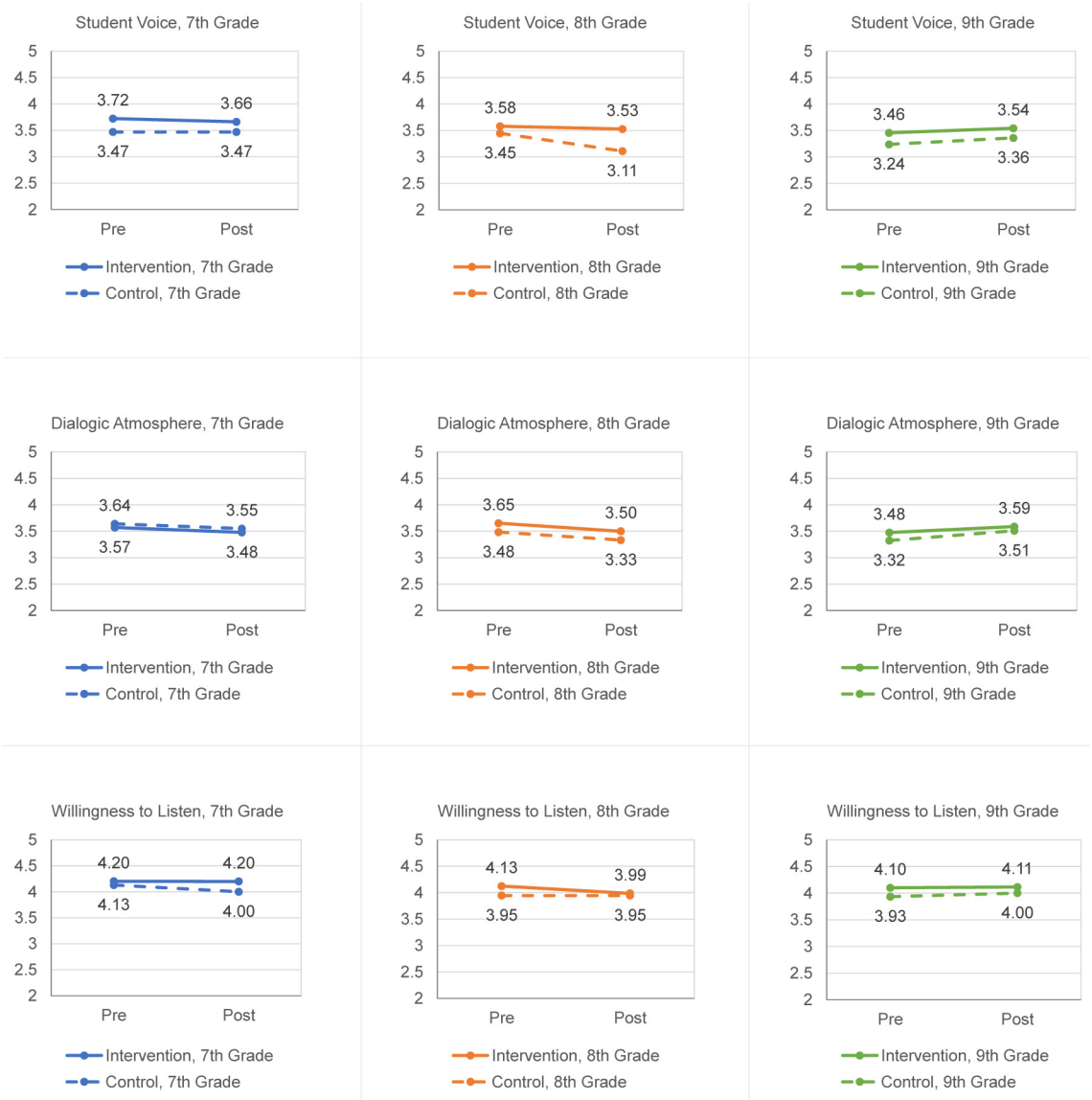


Fig. S1: Dialogic climate in the homeroom classes by condition and age group, matched questionnaires (dependent samples)

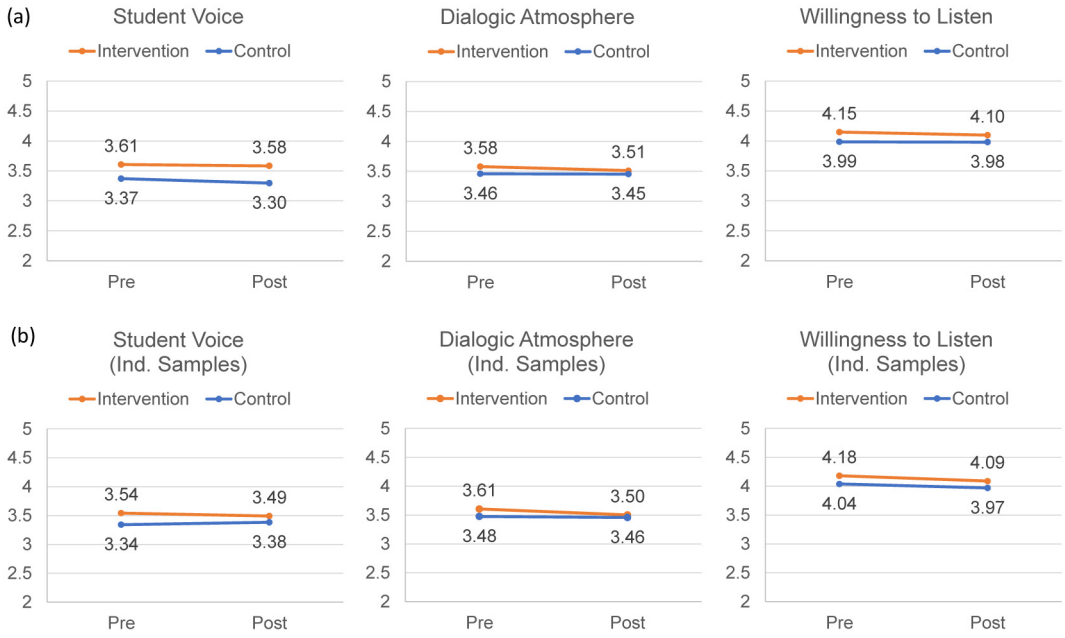


Fig. S2: Dialogic climate in the homeroom classes over time, by time and condition. (a) Matched questionnaires only (dependent samples). (b) DCI questionnaires from all schools except Nofim (independent samples)

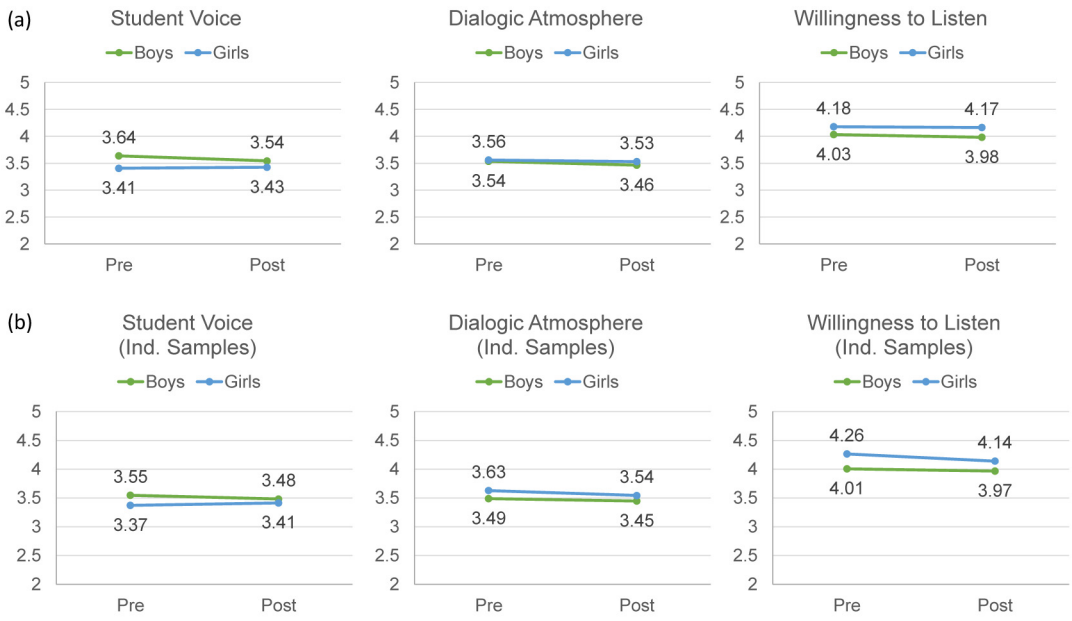


Fig. S3: Dialogic climate in the homeroom classes over time, by time and gender. (a) Matched questionnaires (dependent samples). (b) All DCI questionnaires except Nofim (independent samples)