frontiers in Education

TYPE Original Research
PUBLISHED 18 November 2026
DOI 10.3389/feduc.2026.1982806

Check for updates

OPEN ACCESS

EDITED BY
Mohamed Elsadek,
Tongji University, China

REVIEWED BY
Melyana Nurul Widyawati,
Poltekkes Kemenkes Semarang, Indonesia
Ugur Oner,
Batman University, Türkiye

*CORRESPONDENCE
Alla Prus
✉ prus@zu.edu.ua

RECEIVED 30 August 2026
REVISED 28 October 2026
ACCEPTED 02 November 2026
PUBLISHED 18 November 2026

CITATION
Prus A, Matiash O, Kolomiiets T, Shvets V and Fonariuk O (2026) Repeated passive aromatherapy in mathematics classrooms: no within-lesson psycho-emotional change and challenges of sustained olfactory exposure. Front. Educ. 11:1982806. doi: 10.3389/feduc.2026.1982806

COPYRIGHT
© 2026 Prus, Matiash, Kolomiiets, Shvets and Fonariuk. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Repeated passive aromatherapy in mathematics classrooms: no within-lesson psycho-emotional change and challenges of sustained olfactory exposure

Alla Prus[1]*, Olha Matiash[2], Tamila Kolomiiets[1], Vasyl Shvets[3] and Olena Fonariuk[1]

[1]Department of Mathematics, Zhytomyr Ivan Franko State University, Zhytomyr, Ukraine, [2]Department of Algebra and Methods of Teaching Mathematics, Vinnytsia Mykhailo Kotsiubynskyi State Pedagogical University, Vinnytsia, Ukraine, [3]Department of Mathematics Teaching Methodology, Mykhailo Drahomanov Ukrainian State University, Kyiv, Ukraine

This pilot study investigated psycho-emotional and academic correlates of repeated passive aromatherapy across several weeks of mathematics lessons, and examined whether reduced odour perception limits this delivery mode. A quasi-experimental design with non-equivalent control groups was conducted with 210 school and university participants from three Ukrainian institutions; experimental classes received passive diffusion of eucalyptus and rosemary oils over 2–7 weeks (with an additional lavender phase for university students), whereas control classes received no aromatherapy. Psycho-emotional state was assessed at the beginning and end of lessons, academic performance from grade records, and subjective impressions via questionnaires. Within-group changes were analysed using paired-samples t-tests, and Group × Time interactions — the primary test of a differential intervention effect — were estimated as difference-in-differences (DiD) with Hedges' g. No psycho-emotional subscale showed significant within-lesson change (all $p > .10$), and no significant Group × Time interaction was found for psycho-emotional state (testable only at Zhytomyr) or for grades at Zhytomyr; because psycho-emotional state was measured within lessons, these data do not address cumulative effects of multi-week exposure. At Zvyahel, where the only available control class was a Grade 7 cohort following a different curriculum, grades increased more in the control class than in the experimental classes [DiD = −0.37 grade points, 95% CI (−0.60, −0.14)]; excluding the Grade 5 class reduced this difference [DiD = −0.23, 95% CI (−0.46, −0.003)], which then did not survive multiplicity correction. Given non-random class assignment and curricular differences, this result does not permit causal conclusions in either direction. Despite repeated dose escalation (up to nearly fivefold), half of the questionnaire respondents reported that the scent faded or became imperceptible during sessions and a further 30% were uncertain. This pattern is consistent with olfactory habituation; however, because odour perception was assessed only through self-report and airborne concentration was not measured, habituation cannot be distinguished from an insufficient dose. Attrition and small samples limited

statistical power. Sustained passive diffusion therefore appears difficult to maintain as a reliable classroom intervention. Future studies should evaluate shorter sessions, active inhalation, and intermittent delivery, alongside objective measures of odour perception.

KEYWORDS

aromatherapy, essential oils, intervention fidelity, mathematics classrooms, null effects, olfactory habituation, pilot study

1 Introduction

Mathematics remains one of the most challenging school subjects and, simultaneously, one of the most important for developing analytical thinking, problem-solving abilities, and adaptability in an era of rapid technological change (OECD, 2024; Wang et al., 2025). Beyond its disciplinary content, mathematics is widely regarded as a cognitively demanding domain that engages a broad range of higher-order processes, including working memory, attentional control, logical reasoning, and abstract thinking (Amland et al., 2025). Because successful mathematical performance depends heavily on these cognitive functions, interventions capable of influencing cognition or emotional regulation may plausibly affect learning outcomes.

One underexplored avenue for cognitive enhancement involves the olfactory system. Unlike visual and auditory pathways, olfactory afferent signals bypass the thalamus, projecting first to the olfactory bulb and subsequently to the piriform cortex, amygdala, and entorhinal cortex—structures with direct anatomical connections to the hippocampus and prefrontal cortex (Kostka and Bitzenhofer, 2022). This relatively short subcortical route places olfaction in close proximity to neural systems involved in memory formation, emotional processing, and executive functioning. In addition, nasal airflow has been shown to synchronize neural activity through beta oscillations, potentially influencing cognitive readiness and information processing (Kostka and Bitzenhofer, 2022).

Experimental evidence suggests that essential oils may affect both cognition and emotional functioning. Inhalation of rosemary (Salvia rosmarinus, syn. Rosmarinus officinalis) essential oil improved memory performance in adults under laboratory conditions (Moss et al., 2003), while plasma concentrations of 1,8-cineole—the primary bioactive constituent of rosemary oil—were positively associated with cognitive test performance (Moss and Oliver, 2012). A systematic review of olfactory training — repeated, active sniffing of sets of odours — concluded that such training may support cognitive functioning (Vance et al., 2024), although this intervention differs substantially from the ambient exposure examined here. Beyond direct cognitive effects, an indirect pathway has also been proposed: orange and lavender aromas have been associated with reduced anxiety and improved mood in clinical settings (Lehrner et al., 2000; Lehrner et al., 2005), and pleasant ambient scents have been linked to enhanced task performance through positive affective states (Baron and Thomley, 1994). In student populations, inhalation aromatherapy has been associated with reduced test anxiety (Luan et al., 2023; McCaffrey et al., 2009) and reduced perceived stress (Öner et al., 2026), and a meta-analysis reported anxiolytic effects of lavender, although most included trials were judged to be of low quality (Donelli et al., 2019).

However, evidence regarding the cognitive benefits of aromatherapy remains inconsistent. While some studies have reported improvements in memory, alertness, and mood following exposure to essential oils (Johnson, 2011; Moss et al., 2025), others have found weak, task-dependent, or non-significant effects, particularly for complex cognitive activities requiring sustained executive control (Ilmberger et al., 2001; Tong et al., 2025). These discrepancies suggest that olfactory influences may depend on the specific cognitive demands of the task, characteristics of the odorant, and methodological features of the exposure protocol; in one study, performance after inhalation of essential oils did not differ from that in control conditions, whereas performance was related to participants' subjective evaluations of the odours (Ilmberger et al., 2001).

The potential relevance of aromatherapy to mathematics learning is supported by both cognitive and affective mechanisms. Mathematical competence relies on the coordinated functioning of numerical, linguistic, and spatial processes, with working memory and attentional control consistently identified as major predictors of achievement (Amland et al., 2025; Iglesias-Sarmiento and Deaño, 2011). At the same time, mathematical performance is sensitive to broader affective and motivational states. According to attentional control theory, anxiety-related thoughts can compete with task-relevant processing and consume limited working memory resources, thereby reducing cognitive efficiency (Chang et al., 2026; Eysenck et al., 2007); more generally, negative affect has been associated with reduced cognitive efficiency across a range of tasks. Consequently, interventions capable of improving mood or reducing negative affect may indirectly influence mathematics performance through their effects on emotional functioning.

Despite these theoretical foundations, classroom-based evidence remains very limited. Very few studies have examined aromatherapy in authentic educational settings with school-aged children; a notable exception is Moss et al. (2017), who reported improvements in working memory among children aged 9–11 years following a brief 15 min rosemary exposure. Existing evidence, including studies with student populations, is based predominantly on brief, single-session exposures, often delivered in laboratory settings or in anticipation of a stressful event such as an examination, which limits its generalisability to routine classroom environments. Extended multi-week interventions, comparisons across different essential oils, and investigations using academic performance as an outcome remain scarce.

A particularly underexamined methodological challenge for classroom protocols is olfactory habituation: repeated or

prolonged exposure to an odour typically results in substantial attenuation of perceptual sensitivity within minutes (Dalton, 2000; Jellinek, 1997; Pellegrino et al., 2017). Under continuous passive diffusion conditions—as used in most naturalistic classroom settings—habituation may render the olfactory stimulus effectively absent for much of the lesson, potentially undermining the theoretical mechanism of the intervention. This possibility has rarely been empirically documented in educational contexts.

An additional gap concerns educational environments affected by armed conflict. Chronic stress is associated with impairments in attention, executive functioning, and emotional regulation, all of which are directly relevant to academic performance (Arnsten, 2009). Preliminary evidence also suggests that prolonged stress may alter olfactory processing itself (Shin et al., 2023). Consequently, it remains unclear whether findings obtained under relatively stable laboratory or school conditions can be generalised to educational settings characterised by ongoing wartime stress and disruption. The present study was conducted in Zhytomyr Oblast, Ukraine—a region subject to recurring air alerts, infrastructure damage, and wartime-related psychological stress throughout the study period—providing a naturalistic test of aromatherapy protocols under conditions rarely examined in the research literature.

The present pilot study addresses this gap. Its contribution is threefold: (1) it is, to our knowledge, one of the first multi-week classroom studies of passive aromatherapy to use both psycho-emotional state and routine academic grades as outcomes; (2) it documents the fidelity problem of prolonged passive diffusion — the mismatch between intended and perceived exposure — in authentic lessons; and (3) it was conducted under wartime conditions, which are rarely represented in this literature. We examined whether passive inhalation of rosemary and eucalyptus essential oils during mathematics lessons — with an additional lavender phase in the university sample — was associated with within-lesson changes in psycho-emotional state (SAN subscales: Well-being, Activity, and Mood) and academic performance, and whether it was accompanied by indicators of reduced odour perception.

Given that individual differences in personality may be relevant to participants' psycho-emotional responses to classroom interventions, personality characteristics were also assessed descriptively using the Eysenck Personality Inventory (EPI). These data were used to characterise the study sample rather than to test a separate hypothesis.

On the basis of this theoretical framework, three research questions guided the present study:

RQ1 (psycho-emotional state): Building on evidence that ambient essential oil exposure is associated with reduced anxiety and improved mood (Lehrner et al., 2000; Lehrner et al., 2005), we hypothesized that participation in passive aromatherapy sessions during mathematics lessons would be associated with improvements in psycho-emotional state from the beginning to the end of a lesson, as measured by the SAN subscales (Well-being, Activity, Mood). Because the SAN was administered within lessons and averaged across sessions, RQ1 concerns within-lesson change rather than the cumulative effect of multi-week exposure.

RQ2 (academic performance): Given the proposed direct olfactory–limbic pathway and the indirect affect-mediated pathway linking olfactory stimulation to cognitive functioning (Moss et al., 2003; Moss and Oliver, 2012), we hypothesized that participation in passive aromatherapy sessions would be associated with improved academic performance in mathematics.

RQ3 (intervention fidelity, exploratory): Because olfactory habituation is a well-documented consequence of repeated or prolonged odour exposure (Dalton, 2000; Pellegrino et al., 2017), we explored whether sustained passive diffusion in an authentic multi-week classroom protocol was accompanied by indicators of reduced odour perception, and considered the implications for intervention fidelity.

2 Methods

2.1 Research design

The study employed a quasi-experimental pre–post design with non-equivalent control groups. Three sites participated: Lyceum No. 1 (Zvyahel), Lyceum No. 28 (Zhytomyr), and Zhytomyr Ivan Franko State University (ZhSU). Control groups were included at both school sites: at Lyceum No. 1, class 7B received no aromatherapy intervention; at Lyceum No. 28, class 8V received no aromatherapy intervention. At ZhSU, no control group was available, as only one group of pre-service mathematics teachers existed within the faculty. At Lyceum No. 1, class 7B served as the control class because several parents in this class did not consent to aromatherapy exposure; to avoid exposing pupils without consent, the whole class was assigned to the control condition. Assignment was therefore neither random nor selected by the researchers, and the control class differed from the experimental classes in grade level (Grade 7 vs. Grade 8) and curriculum.

2.2 Participants

A total of 210 participants were enrolled. Lyceum No. 1 (Zvyahel): 115 pupils (88 experimental — classes 8A, 8B, 8V, and 5A; 27 control — class 7B), aged 10–14 years. Lyceum No. 28 (Zhytomyr): 84 pupils (55 experimental — classes 8D and 8E; 29 control — class 8V), aged 13–14 years. ZhSU: 11 pre-service mathematics teachers, aged 19–20 years (no control group). All participants provided informed assent (pupils) or consent (university students); parental consent was obtained for all minor participants.

The study was conducted within the framework of the departmental research project "Professional Competence of the Mathematics Teacher" (UkrINTEI registration No. 0123U100258), approved on 17 February 2025 by the Academic Council of the Faculty of Physics and Mathematics, Zhytomyr Ivan Franko State University. Participation of Lyceum No. 1 named after Lesia Ukrainka (Zvyahel) was formalised through a cooperation agreement with the university (Agreement No. 326, dated 30 December 2023, valid until 31 December 2028). Participation of Lyceum No. 28 named after Hetman Ivan

Vyhovsky (Zhytomyr) was formalised through a separate cooperation agreement (Agreement No. 181, dated 4 November 2022, valid until 31 December 2027). The home institution does not have a dedicated ethics committee; approval by the Academic Council of the relevant faculty is the standard and sufficient procedure for this type of research, and this approval was obtained as described above.

2.3 Intervention

Passive aromatherapy was delivered via open-vessel diffusion of essential oils during mathematics lessons. The core protocol comprised two sequential phases using oils with documented cognitive effects in prior research: Phase 1 used eucalyptus essential oil (Eucalyptus globulus) and Phase 2 used rosemary essential oil (Rosmarinus officinalis L., syn. Salvia rosmarinus Spenn.). Initial dosing was guided by the recommendation of an external consultant with expertise in essential oil pharmacology (typically 5–10 drops per aroma lamp filled with water per 15 m^2 of room area); the consultant recommended an initial dose at the lower end of this range, such that the scent would be barely perceptible. Teachers calculated doses according to their classroom dimensions, adjusting progressively in response to participant reports that the odour became progressively less noticeable during lessons (see Section 3.5 and Table 1). The protocol was adapted across sites in response to participant reactions, logistical constraints, and safety considerations, as described below.

Lyceum No. 1 (Zvyahel). Phase 1 (eucalyptus) ran from 31 March to 22 April 2025 (approximately four weeks, three sessions per week, 35–40 min per session). Phase 2 (rosemary) ran from 23 April to 15 May 2025 (approximately three weeks, three to four sessions per week, 25–35 min per session). Classroom areas ranged from 30 to 42 m^2 with a ceiling height of 3.5 m. Initial dosing in classes 8A and 8B began at 6 drops, increasing to 8–12 drops after approximately one week in response to reduced odour perception, with further weekly increases and smaller increments (2–3 drops) during the rosemary phase. In classes 8 V and 5A, initial dosing began at 4 drops, increasing to 6–10 drops (8V) and 6–8 drops (5A) over the course of the intervention. The control class (7B) received no aromatherapy throughout.

Lyceum No. 28 (Zhytomyr). The intervention ran from 14 April to 9 May 2025. Due to an administrative error, the order of oils at this site was reversed relative to Lyceum No. 1: Phase 1 used rosemary (two weeks, four sessions per week) and Phase 2 used eucalyptus. This limits direct cross-site comparison of phase-specific effects. During the rosemary phase, class 8E maintained a stable dose of 7 drops throughout, whereas in class 8D the dose was reduced (7→4→5→7 drops) during week 1 following reports of headache and dizziness, then held stable at 7 drops in week 2. In class 8E, eucalyptus diffusion continued for two weeks, with the dose increased from 7 to 8 drops in week 2; a further increase to 9 drops was attempted but reverted to 7 drops within 10 min as the scent was judged too strong. In class 8D, eucalyptus diffusion was discontinued after a single 20-minute session on 28 April 2025, following adverse reactions (headache, dizziness) reported by five or more pupils. On the recommendation of the external consultant, eucalyptus was replaced with sweet orange oil (Citrus sinensis) for the remainder of the intervention period in class 8D (30 April–9 May 2025, seven sessions, 7 drops per session). Classroom dimensions were 6 × 7 m with a ceiling height of 3 m. The control class (8V) received no aromatherapy throughout.

Zhytomyr Ivan Franko State University (ZhSU). The university cohort followed a three-phase protocol, as lavender was reserved for the adult sample due to precautionary safety considerations and the lack of prior classroom experience with this oil among younger participants. Phase 1 (eucalyptus) comprised eight sessions from 24 March to 2 April 2025. Phase 2 (rosemary) comprised nine sessions from 7 April to 12 May 2025, with an interruption due to students' teaching practicum. Phase 3 (lavender, Lavandula angustifolia) comprised six sessions from 13 to 21 May 2025. All sessions were 80 min in duration. Classroom areas ranged from 46 to 126 m^2 with a ceiling height of 4 m.

2.3.1 Essential oils

Four commercially available, undiluted (100%) essential oils were used. Eucalyptus (*Eucalyptus globulus* Labill.; organic, steam-distilled; origin: Portugal) and rosemary (*Rosmarinus officinalis* L., syn. *Salvia rosmarinus* Spenn., chemotype cineoliferum) oils were produced by Aroma Fields (Ukraine; ISO 9001:2015-certified producer; technical specification TU U 20.4-3348004730-001:2023). Lavender (*Lavandula angustifolia* Mill., syn. *L. officinalis* Chaix; steam-distilled; origin: France) and sweet orange (*Citrus sinensis* L.; origin: Brazil) oils were produced by Elixan Aromatica GmbH (Switzerland) and distributed under the Vivasan brand. All oils were used within the manufacturers' stated shelf life. The cineole chemotype of rosemary was consistent with the theoretical rationale of the study, which draws on evidence linking 1,8-cineole to cognitive performance (Moss and Oliver, 2012). No independent compositional (e.g., GC–MS) analysis was performed by the authors.

2.3.2 Diffusion procedure

Identical ceramic aroma lamps heated by tea-light candles (metal-cased, approximately 3.5 cm in diameter and 1 cm in height) were used at all sites, with a distance of approximately 10 cm between the flame and the water bowl. At the start of each session, the teacher filled the bowl with 50 mL of water, lit the candle and added the essential oil; the lamp was placed as close to the centre of the room as the classroom layout permitted and remained under the teacher's continuous supervision. The water never evaporated completely, even during 80-minute university sessions, and was discarded after each session. Rooms were ventilated before and after each session but never during it, so that each session began without residual odour from previous exposures. Teachers were instructed to reduce the dose or stop diffusion immediately if any pupil reported discomfort such as headache or dizziness; pupils' well-being took priority over adherence to the protocol (see the dose reduction and discontinuation in class 8D, Section 2.3). Doses were calculated according to room area (Section 2.3; Table 1).

TABLE 1 Essential oil dosing log across study sites.

Institution	Oil	Period	Start (drops)	Final (drops)	Increase	Notes
Lyceum No. 1 (8A)	Eucalyptus	Weeks 1→4	6	8–12	×1.3–2.0	Teacher: 'pupils stopped perceiving the scent' after ~1 week
Lyceum No. 1 (8A)	Rosemary	Weeks 1→3	8–12	progressive, no fixed endpoint	+2–3/lesson	Continuous small increments each lesson
Lyceum No. 1 (8B)	Eucalyptus	Weeks 1→4	6	8–12	×1.3–2.0	Teacher: 'pupils stopped perceiving the scent' after ~1 week
Lyceum No. 1 (8B)	Rosemary	Weeks 1→3	8–12	progressive, no fixed endpoint	+2–3/lesson	Continuous small increments each lesson
Lyceum No. 1 (8 V)	Eucalyptus+Rosemary	Weeks 1→7	4	6–10	×1.5–2.5	Combined trajectory across both oils; not separately reported by phase in teacher log
Lyceum No. 1 (5A)	Eucalyptus + Rosemary	Weeks 1→7	4	6–8	×1.5–2.0	Combined trajectory across both oils; not separately reported by phase in teacher log
University (ZhSU)	Eucalyptus	Sessions 1→8	10	15	×1.5	'Barely perceptible' noted 02.04
University (ZhSU)	Rosemary	Sessions 1→9	7	14	×2.0	'Barely perceptible' noted 05.05
University (ZhSU)	Lavender (a)	Sessions 1→6	7	33	×4.7	Scent undetectable by 19.05 — most extreme case
Lyceum No. 28 (8D)	Rosemary	2 weeks	7	7	Reduced to 4, then restored	Dose reduced mid-week 1 (7→4→5→7) following reports of headache/dizziness; restored to 7 and held stable in week 2
Lyceum No. 28 (8D)	Eucalyptus (b)	Day 1	7	—	—	Protocol discontinued: 5+ adverse reactions within 20 min
Lyceum No. 28 (8D)	Sweet orange	30 Apr–9 May	7	7	None	Introduced after eucalyptus discontinuation; 7 sessions
Lyceum No. 28 (8E)	Rosemary	2 weeks	7	7	None	Stable dose throughout
Lyceum No. 28 (8E)	Eucalyptus	2 weeks	7	8	×1.1 (9 attempted, reverted)	Week 2: increased 7→8→9; the 9-drop dose was reverted to 7 within 10 min as 'too strong'

(a) Most extreme case of reported reduction in perceived scent intensity: dose increased ×4.7 without restoring stable odour perception. (b) Protocol discontinued due to adverse reactions; sweet orange oil subsequently introduced in class 8D. Dose adjustments reflect teacher/supervisor-reported weak or absent scent perception by participants. Lyceum No. 28 control class (8V) received no aromatherapy. Classroom dimensions remained constant for each school class throughout the study; reported dose increases reflect escalation relative to each class's own initial dose, not changes in room size. At ZhSU, by contrast, sessions were held in rooms of varying size (46–126 m^2), and the starting dose for each session was recalculated according to room area using the formula described in Section 2.3; the reported increases for ZhSU therefore reflect escalation beyond this room-adjusted baseline dose. At Lyceum No. 28, the supervising teacher reported, in general terms and without reference to individual identities (the questionnaire responses were anonymous), that complaints of headache or dizziness recurred among a small number of the same pupils across sessions, while most participants reported not perceiving the scent at all, consistent with the anonymous questionnaire responses; the teacher also noted that these pupils were, in her assessment, generally prone to reporting discomfort. This observation is anecdotal and could not be independently verified against the anonymous questionnaire data; see Section 4.3 for a discussion of expectancy (nocebo) effects.

For reference, the lavender and sweet orange oil labels suggest approximately 3 drops per 10 m^2 for room diffusion; no comparable dosing guidance was provided for the eucalyptus and rosemary oils.

2.4 Measures

2.4.1 Psycho-emotional state (SAN)

The SAN questionnaire (Danylchenko and Vertel, 2012; Doskin et al., 1973; Spivak and Osmanova, 2023) is a validated instrument assessing three dimensions of psycho-emotional state: Well-being (W), Activity (A), and Mood (M), via 30 bipolar adjective pairs. In the original version, items are rated on a 7-point semantic differential scale; in the present study, a 5-point response format was used to reduce cognitive load for school-aged participants. Each subscale comprises 10 items, yielding scores ranging from 10 to 50, with higher scores indicating a better psycho-emotional state. Because the Ukrainian wording of some items in the version used (e.g., item 13 "tense—relaxed" and item 16 "calm—agitated") differs from the original instrument, items were scored according to the semantic direction of their poles, and this scoring was verified empirically through inter-item correlations: items 3, 4, 9, 10, 13, 15, 21, 22, 27, and 28 (positive pole on the right) were scored

directly, and the remaining items were reverse-scored (6−x), so that higher scores indicate a better psycho-emotional state. Internal consistency in the present sample was high (Cronbach's $\alpha = .92$, .89, and .95 for Well-being, Activity, and Mood, respectively; 922 school questionnaires). The full item list and scoring key are provided in Supplementary Table S1. The SAN was administered at the beginning and end of each lesson session in which it was used. Due to a technical omission in the online questionnaire form used with the university sample, one Activity-subscale item (item 16, "calm—agitated") was inadvertently excluded; Activity scores for this group are therefore based on nine items and were pro-rated to a 10-item equivalent (×10/9). A simulation in the school data, where all ten items were available, showed that omitting this item and pro-rating lowered individual Activity scores by 0.63 points on average ($r = .992$ with the full score), with individual differences ranging from −4.0 to +1.8 points; however, participant-level pre–post change scores were almost unaffected ($r = .97$–$.99$ between change scores computed with and without the item), and paired-test results were essentially unchanged (e.g., $p = .199$ vs.187 at Lyceum No. 1).

2.4.2 Academic performance

Mathematics grades were recorded from the school electronic gradebook system ("Novi Znannia" platform). Only current grades were used, i.e., all grades recorded for individual lessons, including grades for independent work, tests, and notebook checks; thematic, semester, and annual grades were excluded, as they represent derived aggregates of lesson grades rather than individual assessment events. At both sites, PRE grades were lesson grades recorded from the start of the second semester (early January 2025) until the onset of the intervention at that site, and POST grades were lesson grades recorded during the intervention period itself. This definition was applied consistently across all classes at each site, including the control classes, enabling individual-level paired pre–post comparison.

2.4.3 Subjective impressions

Participants' subjective responses to the aromatherapy intervention were assessed using two purpose-designed questionnaires administered at different points in the study.

The post-transition questionnaire was administered at the start of the first session of the second oil phase, prior to exposure to the new scent. It assessed participants' retrospective experience with the first oil, including: perceived olfactory habituation (whether the scent had faded during sessions), responses to dose escalation, and perceived cognitive effects. One item on perceived cognitive effects was preceded by a brief statement of prior research findings regarding the first oil; responses to this item may therefore reflect demand characteristics and should be interpreted with caution.

The end-of-study questionnaire was administered at the final session of the intervention. It assessed: self-reported changes in concentration ability during lessons, emotional responses to the scents used throughout the study, comfort levels during lessons, willingness to participate in future aromatherapy research, and overall interest in the experiment (rated on a 1–5 scale).

2.4.4 Personality (EPI)

The Eysenck Personality Inventory (EPI) was administered once to assess Extraversion (E) and Neuroticism (N) (Eysenck and Eysenck, 1964). The Lie scale (L) was included to assess response validity; scores of $L \geq 5$ were considered indicative of potentially invalid responses, consistent with standard EPI scoring conventions.

2.5 Procedure

Data collection proceeded according to the phase structure described in Section 2.3. In each phase, the SAN was administered at the start and end of each session; lesson grades were recorded from the electronic gradebook throughout. The post-transition questionnaire was administered at the beginning of the first session of the second oil phase, and the end-of-study questionnaire at the final session. The EPI was administered once during the study period, at a session designated by the supervising teacher or instructor. Control classes received no aromatherapy. The control class at Lyceum No. 28 (8V) completed the SAN questionnaire on the same schedule as the experimental groups, whereas the questionnaire was not administered consistently in the control class at Lyceum No. 1 (7B; see Section 3.1).

2.6 Data analysis

At each lesson, participants completed the SAN questionnaire immediately before aroma exposure and again at the end of the lesson. Because multiple lesson-level observations were obtained from each participant, all pre-lesson scores were averaged to obtain a single pre-lesson value, and all post-lesson scores were averaged to obtain a single post-lesson value for each participant. Statistical analyses were then conducted at the participant level to avoid pseudoreplication and to preserve the individual as the unit of analysis. SAN responses were collected via Google Forms, which recorded each respondent's e-mail address; addresses were used only to link a participant's responses across sessions and were removed after linkage. Inclusion in the matched analyses required at least one pre-lesson and one post-lesson response; because attendance varied, pre- and post-lesson averages were not always based on the same sessions.

The number of matched responses was lower than the number of enrolled participants because some questionnaires were not completed due to lesson interruptions, student absences, and occasional missing responses. Paired t-tests were used to assess pre–post differences within each group. Cohen's d was calculated as a measure of effect size (Cohen, 1988). A sensitivity power analysis (two-tailed $\alpha = .05$, power = .80) was conducted to determine the minimum effect sizes detectable with the achieved matched samples.

For mathematics grades, paired t-tests were used at both sites for all classes, including control classes. At Lyceum No. 28, individual lesson grades from the start of the second semester to

13 April 2025 served as PRE values, and individual lesson grades during the intervention period (14 April—9 May 2025) as POST values. At Lyceum No. 1, individual lesson grades from the second semester prior to 31 March 2025 served as PRE values, and individual lesson grades during the intervention period (31 March—15 May 2025) as POST values, for both experimental classes and the control class (7B). Grade 5A covers a different curriculum from Grade 8 classes and is reported separately. EPI data were examined descriptively; no inferential statistics were applied to personality variables. For these secondary paired comparisons of grades, Holm corrections were applied within each site's family of tests (Tables 2, 3); the secondary paired SAN comparisons (Table 4) are reported without correction, consistent with their exploratory status.

In addition to the within-group paired-samples comparisons reported above, a Group × Time interaction was estimated for each outcome and site to test directly whether the pre–post change differed between experimental and control conditions. For designs with one pre and one post measurement per participant, this interaction is mathematically equivalent to an independent-samples comparison of within-participant change scores (Δ = post−pre); we report this difference-in-differences (DiD) estimate with its 95% confidence interval and Hedges' g, cross-validated with a mixed-design ANOVA (Group × Time). Welch's *t*-test was used throughout to accommodate unequal group variances.

This interaction analysis was possible only where a genuine control condition existed with matched pre- and post-measurements. For the SAN questionnaire, the Zvyahel control class (7B) had near-total non-participation in the survey (see §4.6) and could not support a between-group comparison; the interaction was therefore estimated only at the Zhytomyr site (Lyceum No. 28, class 8V control vs. 8E/8D experimental). For academic grades, individual-level pre-intervention data were reconstructed from grade-book records predating the intervention onset (31 March 2025 at Zvyahel; 14 April 2025 at

TABLE 2 Mathematics Grades — Pre/Post Analysis, Zhytomyr (Lyceum No. 28).

Class (Role)	Subject	*n*	PRE M (SD)	POST M (SD)	Δ	t	*p*	d [95% CI]
8V (Control)	Algebra	29	5.76 (1.57)	6.16 (1.65)	+0.40	2.264	.032*	0.42 [0.04, 0.80]
8V (Control)	Geometry	29	6.01 (1.65)	5.77 (1.74)	−0.24	−1.568	.128	0.29 [−0.08, 0.66]
8E (Exp)	Algebra	27	6.09 (1.14)	6.43 (1.28)	+0.35	2.649	.014*	0.51 [0.11, 0.91]
8E (Exp)	Geometry	27	5.74 (1.33)	5.75 (1.31)	+0.01	0.072	.944	0.01 [−0.36, 0.39]
8D (Exp)	Algebra	28	5.59 (1.72)	5.64 (1.86)	+0.05	0.393	.698	0.07 [−0.30, 0.45]
8D (Exp)	Geometry	28	5.75 (1.81)	5.62 (1.82)	−0.14	−1.309	.202	0.25 [−0.13, 0.62]

PRE = individual student mean of all current grades in Semester 2 before the intervention (before 14 April 2025); POST = individual student mean of current grades during the intervention (14 April–9 May 2025). Current grades include grades for independent work, tests, and notebook checks; thematic, semester, and annual grades were excluded. Values were recomputed during revision from the class gradebook records using the same rule as for Lyceum No. 1 (Table 3). *n* = pupils with at least one grade in the subject in both periods. d = Cohen's d (absolute value); interpreted following Cohen (1988). *$p < .05$ (uncorrected). Two results were nominally significant before correction (8E Algebra, $p = .014$; 8V Algebra, control class, $p = .032$); neither remained significant after Holm correction for this six-test family ($p_{Holm} = .081$ and .158).

TABLE 3 Mathematics Grades — Pre/Post Analysis, Zvyahel (Lyceum No. 1).

Class (Role)	Subject	*n*	PRE M (SD)	POST M (SD)	Δ	t	*p*	d [95% CI]
Grade 8 Classes — Experimental (Algebra & Geometry)								
8A (Exp)	Algebra	23	7.87 (1.65)	8.22 (1.54)	+0.35	2.856	.009**	0.60 [0.15, 1.04]
8A (Exp)	Geometry	22	7.89 (1.72)	8.11 (1.87)	+0.23	2.103	.048*	0.45 [0.01, 0.89]
8B (Exp)	Algebra	22	7.01 (2.07)	7.23 (1.94)	+0.22	2.645	.015*	0.56 [0.11, 1.01]
8B (Exp)	Geometry	22	7.19 (2.12)	7.55 (1.98)	+0.36	3.253	.004**	0.69 [0.23, 1.16]
8V (Exp)	Algebra	23	8.17 (1.88)	7.89 (1.96)	−0.28	−1.835	.080	0.38 [−0.04, 0.81]
8V (Exp)	Geometry	23	7.64 (1.79)	8.03 (1.99)	+0.39	2.782	.011*	0.58 [0.14, 1.02]
Grade 5 — Experimental (Mathematics)								
5A (Exp)	Mathematics	20	8.18 (2.41)	7.76 (2.32)	−0.42	−4.213	<.001**	0.94 [0.42, 1.47]
Control Class (Algebra & Geometry)								
7B (Ctrl)	Algebra	27	7.27 (2.29)	7.60 (2.10)	+0.32	2.799	.010**	0.54 [0.14, 0.94]
7B (Ctrl)	Geometry	27	7.08 (2.23)	7.68 (2.01)	+0.60	4.216	<.001**	0.81 [0.38, 1.25]

PRE = individual student mean of all current grades in Semester 2 before the intervention (before 31 March 2025); POST = individual student mean of current grades during the intervention period (31 March–15 May 2025), the same window for all classes. Current grades include grades for independent work, tests, and notebook checks; thematic, semester, and annual grades were excluded. d = Cohen's d (absolute value); interpreted following Cohen (1988). *$p < .05$; **$p < .01$ (uncorrected). After Holm correction for this nine-test family, three results remained significant: 8B Geometry ($p_{Holm} = .027$), 5A Mathematics (a decrease; $p_{Holm} = .004$), and 7B Geometry (control class; $p_{Holm} = .002$). 5A covers a different curriculum (Grade 5 Mathematics). Paired n values include all participating pupils with grade records in both periods; pupils studying remotely had no grade records and were therefore not included.

TABLE 4 SAN Paired *t*-test Results (Student-level Analysis).

Group/Scale	n pairs	M pre (SD)	M post (SD)	Δ	t	*p*	Sig.	d [95% CI]
Lyceum No. 1 — Experimental								
W — Well-being	24	33.2 (7.0)	32.5 (6.2)	−0.8	0.706	.487	n.s.	0.14 [−0.26, 0.54]
A — Activity	24	30.2 (7.2)	28.6 (6.7)	−1.6	1.322	.199	n.s.	0.27 [−0.14, 0.68]
M — Mood	24	35.4 (7.3)	33.2 (6.6)	−2.2	1.681	.106	n.s.	0.34 [−0.07, 0.75]
Lyceum No. 28 — Experimental								
W — Well-being	46	34.6 (7.2)	33.5 (8.1)	−1.2	1.058	.296	n.s.	0.16 [−0.13, 0.45]
A — Activity	46	30.3 (7.1)	31.5 (7.3)	+1.2	−1.283	.206	n.s.	0.19 [−0.10, 0.48]
M — Mood	46	35.7 (7.8)	35.0 (9.1)	−0.8	0.630	.532	n.s.	0.09 [−0.20, 0.38]
Lyceum No. 28 — Control								
W — Well-being	25	37.0 (7.1)	37.4 (7.0)	+0.4	−0.728	.474	n.s.	0.15 [−0.25, 0.54]
A — Activity	25	33.1 (7.0)	33.6 (6.5)	+0.6	−1.265	.218	n.s.	0.25 [−0.15, 0.65]
M — Mood	25	39.2 (6.4)	38.5 (7.1)	−0.7	1.226	.232	n.s.	0.25 [−0.15, 0.64]
University students (ZhSU)								
W — Well-being	11	35.5 (6.1)	36.6 (5.5)	+1.1	−1.724	.115	n.s.	0.52 [−0.11, 1.15]
A — Activity[1]	11	31.0 (4.5)	31.2 (3.7)	+0.2	−0.292	.776	n.s.	0.09 [−0.50, 0.68]
M — Mood	11	36.7 (5.3)	37.4 (5.0)	+0.7	−1.690	.122	n.s.	0.51 [−0.12, 1.14]

Δ = M post−M pre; positive values indicate a more favourable state after the session; negative t indicates post > pre. n.s., not significant. d = Cohen's d (absolute value), calculated as the mean of paired differences divided by the standard deviation of paired differences; interpreted following Cohen (1988): ∼0.2 = small, ∼0.5 = medium, ∼0.8 = large. No corrections for multiple comparisons were applied given the exploratory nature of this analysis; with 12 tests at α = .05, approximately 0.6 significant results would be expected by chance alone, compared with none observed. [1]Pro-rated nine-item score; see Table 5 note.

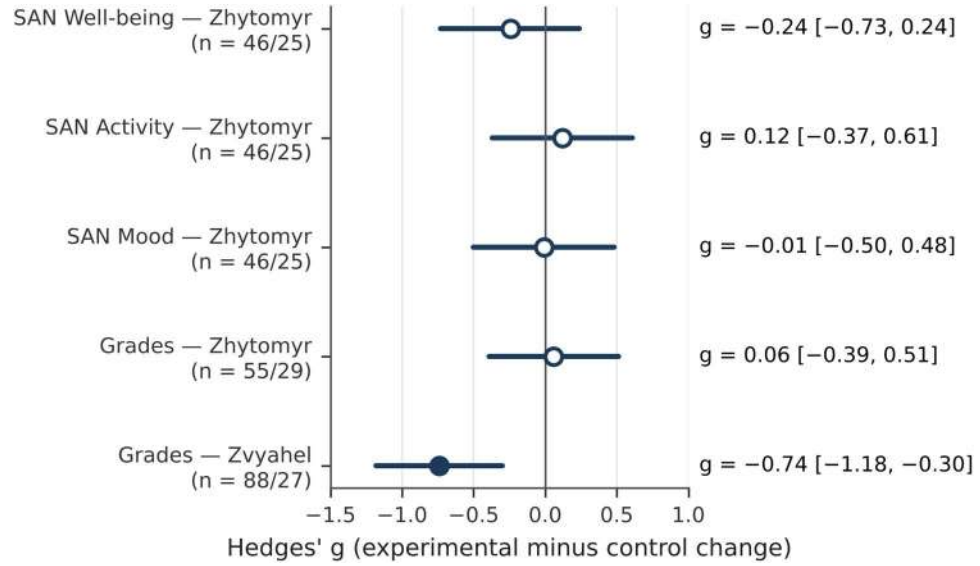


FIGURE 1
Standardised group × time interaction effects (Hedges' g with 95% confidence intervals) for the five primary tests. Negative values indicate a smaller pre–post improvement in the experimental than in the control group. Only the Zvyahel grades effect was significant after Holm correction; see Section 3.4 for the design limitations affecting this comparison.

Zhytomyr) at both sites, allowing the interaction to be tested at both sites.

Given regression to the mean is a known concern whenever groups differ, even non-significantly, at baseline, we additionally fit an ANCOVA (change score regressed on baseline level and group) for the Zvyahel grades result to determine whether the observed interaction persisted after adjusting for baseline differences. During revision, sensitivity analyses examined (a) exclusion of the Grade 5 class, (b) equal weighting of algebra and geometry, and (c) a shorter post-intervention window (31

March—30 April) (Supplementary Table S2). During revision, statistical computations were re-run from the raw data using Python scripts prepared with the assistance of Claude (Anthropic); the authors reviewed all results and take full responsibility for them. In the second round of revision, the Lyceum No. 28 grade analyses were recomputed from the class gradebook records using the same rule as at Lyceum No. 1, and SAN responses at this site were reassigned to the correct group (Section 3.1).

Five primary Group × Time tests were conducted (three SAN subscales and academic grades at Zhytomyr; academic grades at Zvyahel). *P*-values were Holm-corrected across this full set of five tests, rather than within each outcome table separately, to reflect the true family of primary comparisons. The pre-existing paired-samples t-tests (Tables 2–4) are retained but are reclassified as secondary/exploratory analyses; the Group × Time interaction is the primary analysis addressing the research question of whether the intervention had a differential effect relative to control. This primary family was defined during the first round of revision, in response to reviewer feedback, and is therefore *post hoc*.

3 Results

3.1 Participant characteristics

A total of 210 participants were enrolled across three Ukrainian educational institutions. Lyceum No. 1 (Zvyahel): 115 pupils aged 10–14 years (88 experimental across four classes — 8A, 8B, 8V, and 5A; 27 control — class 7B). Lyceum No. 28 (Zhytomyr): 84 pupils aged 13–14 years (55 experimental across two classes — 8D and 8E; 29 control — class 8V). Zhytomyr Ivan Franko State University: 11 pre-service mathematics teachers aged 19–20 years, participating in an exploratory capacity without a control group.

Not all enrolled participants provided complete data for each measure. Matched pre–post pairs available for SAN analysis were: Lyceum No. 1 experimental — $n = 24$; Lyceum No. 28 experimental — $n = 46$; Lyceum No. 28 control — $n = 25$; university students — $n = 11$. A matched SAN sample for the Lyceum No. 1 control class (7B) is not reported: the supervising teacher understood psycho-emotional assessment to apply only to aromatherapy-exposed classes and consequently did not administer the questionnaire consistently to this class, resulting in an insufficient number of responses for analysis (see Section 4.6). Reduced sample sizes for the remaining groups reflect participant absences due to wartime disruptions and incomplete questionnaire responses. Participants included in the matched samples did not differ significantly in baseline (pre-lesson) SAN scores from those with pre-lesson data only (Lyceum No. 1: all $p > .14$; Lyceum No. 28 experimental: all $p > .72$). Group membership was recorded through a question in the online form. At Lyceum No. 28, four pupils from the control class (8V) selected the experimental-group option in some sessions; during revision, all responses linked to their e-mail addresses were assigned to the control group to which they belonged, and all SAN analyses at this site were recomputed. Responses from these pupils had previously been counted in both groups.

3.2 Psycho-emotional state (SAN questionnaire)

SAN subscale scores were analysed at the student level, with each participant's sessions averaged to yield one pre-score and one post-score. Descriptive statistics are presented in Table 5, including the university student sample.

Paired t-tests revealed no statistically significant pre–post changes on any SAN subscale in any group, including university students (all $p > .10$; all Cohen's $d \leq 0.52$; Table 4). In the school experimental groups, the small non-significant changes were mostly in the direction of a slightly less favourable state at the end of sessions (e.g., Mood at Lyceum No. 1: $\Delta = -2.2$, $p = .106$), whereas university students showed small non-significant improvements in Well-being ($d = 0.52$, $p = .115$) and Mood ($d = 0.51$, $p = .122$). The Activity result for university students was unchanged under alternative item assignments for the university questionnaire form (all $p > .59$).

Reduced matched sample sizes reflect participant absences due to wartime disruptions and incomplete questionnaire responses (Lyceum No. 1 experimental: $n = 24$ pairs; Lyceum No. 28 experimental: $n = 46$ pairs; Lyceum No. 28 control: $n = 25$ pairs; university students: $n = 11$ pairs). Pre-existing differences between experimental and control groups at Lyceum No. 28 at baseline were not analysed inferentially, as the quasi-

TABLE 5 SAN Descriptive Statistics: Mean (SD) by Group.

Group	*n*	W — Well-being M (SD)	A — Activity M (SD)	M — Mood M (SD)
Lyceum No. 1, Zvyahel				
Exp — Pre (all sessions avg.)	44	32.0 (8.2)	28.7 (7.4)	33.7 (8.6)
Exp — Post (all sessions avg.)	32	32.2 (7.3)	28.7 (7.3)	32.9 (8.1)
Lyceum No. 28, Zhytomyr				
Exp — Pre	54	34.6 (7.3)	30.3 (7.3)	35.9 (8.0)
Exp — Post	47	33.5 (8.0)	31.3 (7.4)	35.0 (9.0)
Control — Pre	26	37.3 (7.1)	33.4 (7.2)	39.4 (6.4)
Control — Post	25	37.4 (7.0)	33.6 (6.5)	38.5 (7.1)
University students (ZhSU)				
Exp — Pre (all sessions avg.)	11	35.53 (6.13)	30.99 (4.52)[1]	36.71 (5.26)
Exp — Post (all sessions avg.)	11	36.65 (5.55)	31.23 (3.74)[1]	37.41 (5.01)

Values are student-level means for all participants providing at least one response in the respective phase. Scale range 10–50 for all subscales; higher scores indicate a better psycho-emotional state. W = Well-being; A = Activity; M = Mood. At Lyceum No. 28, group membership was recorded through a question in the online form; four pupils from the control class (8V) selected the experimental-group option in some sessions. During revision, all responses linked to their e-mail addresses were assigned to the control group, and all Lyceum No. 28 SAN analyses (Tables 4–6) were recomputed; responses from these pupils had previously been counted in both groups. [1] Item 16 was missing from the university questionnaire form; Activity scores for this group are pro-rated from nine items (×10/9; see Section 2.4.1).

TABLE 6 Group × Time Interaction Analysis: SAN Subscales and Academic Grades

Site	Outcome	n (exp/ctrl)	DiD	95% CI	p (raw)	p (Holm, $n = 5$)	Hedges' g
Zhytomyr	SAN — Well-being	46 / 25	−1.57	[−4.05, 0.91]	.211	.843	−0.24
Zhytomyr	SAN — Activity	46 / 25	+0.65	[−1.45, 2.75]	.538	1.000	0.12
Zhytomyr	SAN — Mood	46 / 25	−0.07	[−2.71, 2.58]	.959	1.000	−0.01
Zhytomyr	Academic grades	55 / 29	+0.03	[−0.22, 0.28]	.818	1.000	0.06
Zvyahel	Academic grades	88 / 27	−0.37	[−0.60, −0.14]	.0023	.012	−0.74

DiD = difference in mean pre–post change (experimental minus control), in scale points (10–50) for SAN subscales and in grade points (1–12) for academic grades. p (Holm, $n = 5$) is Holm-Bonferroni corrected across all five primary Group×Time interaction tests jointly (see §2.6), not within each outcome row separately. n = participants included in each analysis: matched pre–post pairs for SAN subscales; pupils with current grades in both periods for academic grades. The Zvyahel academic-grades row is the only result significant after correction.

TABLE 7 ANCOVA of Zvyahel Grade Change Score on Baseline Level and Group.

Term	Coefficient	p
Intercept	0.960	<.001
Baseline grade (grade_pre)	−0.074	<.001
Group (experimental = 1)	−0.329	.002

experimental design with non-equivalent groups precludes meaningful between-group comparison (see Section 4.6).

3.3 Academic performance in mathematics

Mathematics grades were analysed separately for each site. At Lyceum No. 28 (Zhytomyr), individual-level pre–post comparisons were conducted using paired t-tests, with the pre-intervention period defined as grades recorded before 14 April 2025 and the post-intervention period as grades recorded during the intervention (14 April—9 May 2025). Thematic and semester grades were excluded from analysis as they are derived aggregates of current grades. During revision, these values were recomputed from the class gradebook records using the same rule as at Lyceum No. 1 (all current grades, including independent work, tests, and notebook checks). Results are presented in Table 2.

Algebra grades increased slightly in one experimental class and in the control class (8E: $\Delta = +0.35$, $p = .014$, $p_{\text{Holm}} = .081$, $d = 0.51$; control class 8V: $\Delta = +0.40$, $p = .032$, $p_{\text{Holm}} = .158$, $d = 0.42$), whereas class 8D showed no change ($\Delta = +0.05$, $p = .698$). No significant pre–post changes were observed for geometry in any class (all $p > .12$), and no result survived Holm correction (Table 2).

At Lyceum No. 1 (Zvyahel), individual-level paired t-tests were conducted using individual pre-intervention grades from the second semester (collected prior to 31 March 2025) as PRE values, and individual lesson grades during the intervention period (31 March—15 May 2025) as POST values. Thematic and semester grades were excluded. Grade 8 classes (8A, 8B, 8V) and the control class (7B) covered algebra and geometry; class 5A covered the mathematics curriculum for that year level, which differs substantially in content and should not be compared directly with Grade 8 results. Results are presented in Table 3.

Among the experimental classes, the Grade 8 classes showed small improvements in most comparisons (8A: algebra $\Delta = +0.35$, geometry $\Delta = +0.23$; 8B: algebra $\Delta = +0.22$, geometry $\Delta = +0.36$; 8V: geometry $\Delta = +0.39$), with the exception of 8V algebra ($\Delta = -0.28$, $p = .080$); only the 8B geometry result survived Holm correction ($p_{\text{Holm}} = .027$). Class 5A showed a decrease in mathematics grades ($\Delta = -0.42$, $p < .001$, $p_{\text{Holm}} = .004$). The control class (7B), which received no aromatherapy, improved in both algebra ($\Delta = +0.32$, $p = .010$, $p_{\text{Holm}} = .055$) and geometry ($\Delta = +0.60$, $p < .001$, $p_{\text{Holm}} = .002$), with effect sizes comparable to or larger than those in the experimental classes (Table 3).

3.4 Group × time interaction analysis (primary analysis)

Because the Zvyahel control class (7B) provided no usable SAN data, Group × Time analyses for psycho-emotional state were possible only at the Zhytomyr site. For academic grades at Zvyahel, the control class was a Grade 7 cohort following a different curriculum from the Grade 8 experimental classes, and class assignment was not random (Section 2.1); interactions at this site therefore cannot be interpreted causally. No significant Group × Time interaction was found for any SAN subscale at the Zhytomyr site [Well-being: DiD = −1.57 points, 95% CI (−4.05, 0.91), $g = -0.24$, $p = .211$; Activity: $g = 0.12$, $p = .538$; Mood: $g = -0.01$, $p = .959$; all Holm-adjusted $p \geq .843$; Table 6; Figure 1]. For academic grades, the Zhytomyr site likewise showed no significant interaction [DiD = +0.03 grade points, 95% CI (−0.22, 0.28), $g = 0.06$, $p = .818$; $n = 55$ experimental and 29 control pupils with grades in both periods].

At the Zvyahel site, the interaction for academic grades was significant and survived Holm correction across the full family of five primary tests [DiD = −0.37 grade points, 95% CI [−0.60, −0.14], Welch $t(41.0) = -3.25$, $p = .0023$, Holm-adjusted $p = .012$; mixed ANOVA $F(1,113) = 11.33$, $p = .001$, $\eta p^2 = .091$; $n = 88$ experimental and 27 control pupils]: the control class (7B) improved more over the study period ($\Delta = +0.43$) than the pooled experimental classes ($\Delta = +0.06$; 5A, 8A, 8B, 8V individually: −0.42, +0.29, +0.28, +0.02). The student-level outcome was the mean of all current grades (algebra and geometry combined; mathematics for 5A) in each period. Baseline (pre-intervention) grades did not differ significantly between conditions (experimental M = 7.76, control M = 7.20,

$p = .254$). Sensitivity analyses (Supplementary Table S2) showed that the direction of the effect was consistent across analytic choices, but its magnitude and significance depended on them: weighting algebra and geometry equally (DiD = −0.40, $p = .0015$) or restricting the post-intervention window to 31 March—30 April (DiD = −0.47, $p < .001$) yielded similar or larger effects, whereas excluding the Grade 5 class reduced the effect [DiD = −0.23, 95% CI (−0.46, −0.003), $p = .048$], which would not survive Holm correction.

Because baseline levels, while not significantly different, favoured the experimental group, we tested whether regression to the mean could account for this interaction. Baseline grade correlated negatively with change score overall ($r = -.32$, $p < .001$), within the control group ($r = -.49$, $p = .009$), and within the experimental group ($r = -.24$, $p = .024$), consistent with a regression-to-the-mean contribution. An ANCOVA regressing the change score on baseline level and group (Table 7) showed that the group effect remained similar in magnitude after this adjustment ($\beta = -0.33$ vs. unadjusted DiD = −0.37, $p = .002$), indicating that regression to the mean alone is unlikely to account for the observed interaction.

3.5 Dose adjustments and reported odour perception

Progressive dose escalation was documented throughout the study at all sites where diffusion was sustained, consistent with reduced olfactory perception over repeated exposure, as reported by teachers and supervisors (Table 1; Figure 2).

At Lyceum No. 1 (Zvyahel), the teacher reported that pupils ceased to perceive the eucalyptus scent within the first week, prompting a doubling of the initial dose (from 6 to 12+ drops) over four weeks. Continuous additions of 2–3 drops per session were similarly required during the rosemary phase. At Zhytomyr Ivan Franko State University, reduced scent perception was reported across all three oils; the most pronounced case involved lavender, where the dose increased from 7 to 33 drops (×4.7) without restoring stable odour perception. At Lyceum No. 28 (Zhytomyr), rosemary dosing remained stable at 7 drops throughout in class 8E; in class 8D, the dose was reduced (7→4→5→7 drops) during week 1 following reports of headache and dizziness, then held stable at 7 drops in week 2. Eucalyptus diffusion in class 8D was discontinued after a single 20-minute session due to five or more students reporting adverse reactions (headache, dizziness). Sweet orange oil (Citrus sinensis) was subsequently introduced in class 8D for seven sessions (30 April—9 May 2025) at a stable dose of 7 drops per session, with no further adverse reactions reported.

3.6 Subjective impressions

Post-transition questionnaire data ($n = 80$; Lyceum No. 28: $n = 35$, Lyceum No. 1: $n = 35$, university: $n = 10$) indicated that half of the respondents (40/80) reported reduced scent perception during sessions (36% noted that the scent faded within the first few minutes of each session; 14% reported that it became entirely imperceptible), and a further 30% were uncertain. Following dose increases, 69% (55/80) reported that the intensified scent again faded shortly thereafter. These

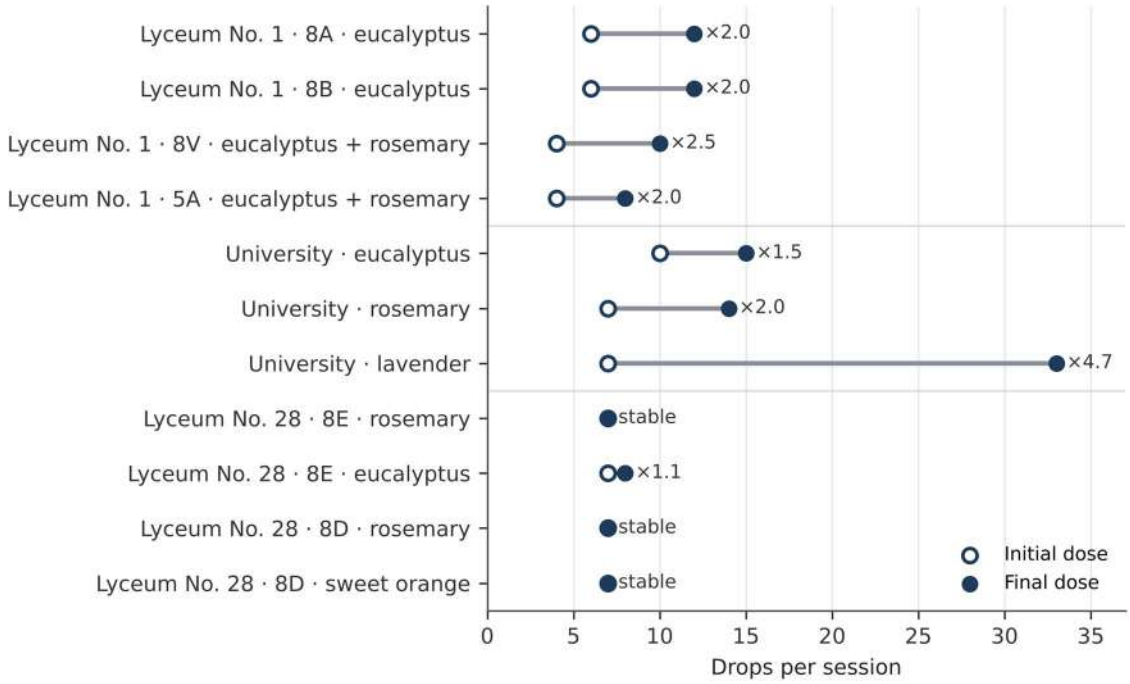


FIGURE 2
Initial and final essential-oil doses (drops per session) by site, class, and oil. Escalation reflects teacher- and supervisor-reported weak or absent scent perception; at Lyceum No. 28, doses were held stable or temporarily reduced because of adverse reactions (see Table 1). The rosemary phase at Lyceum No. 1 (8A, 8B), in which doses were increased by 2–3 drops per lesson without a fixed endpoint, is not shown.

subjective reports are consistent with the dosing patterns presented in Table 1. Perceived cognitive improvement was reported by only 21% (17/80) of participants; however, this item was subject to potential demand characteristics given that participants were first informed of prior research findings before being asked whether they had noticed similar effects, and should therefore be interpreted with caution. Self-reported cognitive benefits were relatively uncommon and were not reflected in objective academic outcomes.

End-of-study questionnaire data ($n = 66$) indicated that 59% (39/66) of participants noticed no change in their ability to concentrate; 23% (15/66) reported a slight improvement; and 12% (8/66) reported a marked improvement; of the remaining four respondents, two reported a slight or marked deterioration, one reported improvement with one scent and slight deterioration with another, and one gave no definite answer. Most participants described the aromas as neutral or mildly pleasant (35% mildly pleasant; 30% neutral); 12% (8/66) reported mild irritation. Adverse reactions were reported by a minority of participants: three individuals (5%) described clearly negative emotional responses including headache or strong discomfort, and 12% (8/66) reported feeling somewhat or completely uncomfortable during lessons. Willingness to participate in future aromatherapy research was expressed by 58% (38/66) of participants. Mean interest rating: 3.65/5.00 (SD = 1.14; $N = 66$); when disaggregated, university students rated interest markedly higher (M = 4.55, $n = 11$) than schoolchildren (M = 3.47, $n = 55$).

Teachers at both school sites independently reported increased pupil activity and willingness to participate during lessons over the course of the intervention. At Lyceum No. 1, the supervising teacher's field notes described an initial "adjustment" phase, during which some pupils expressed curiosity and others apprehension toward the scent, followed by a "working" phase, during which most pupils appeared to habituate to the odour while showing increased attentiveness during lessons.

3.7 Participant characteristics: personality profile

To provide descriptive context for the participant sample, EPI data were examined for Extraversion, Neuroticism, and response validity. EPI data were available for 64 participants (Lyceum No. 1: $n = 13$; Lyceum No. 28: $n = 40$; university students: $n = 11$). Between 32% and 36% of responses were flagged as potentially unreliable ($L \geq 5$), consistent with social desirability tendencies documented in adolescent and young adult samples. Given this limitation, EPI data are reported descriptively only; no inferential statistics were applied. Descriptive statistics for Extraversion, Neuroticism, and the Lie scale, by group, are presented in Table 8.

Among schoolchildren ($n = 53$), mean Extraversion was 12.68 (SD = 3.87), indicating ambiversion, and mean Neuroticism was 13.62 (SD = 5.22), indicating an average level of emotional stability. Among university students ($n = 11$), mean Extraversion was 10.36 (SD = 4.20), also within the ambiversion range, and mean Neuroticism was 17.00 (SD = 3.41), which falls within the range conventionally labelled as emotional instability on this instrument; given the very small sample size, this descriptive observation should not be over-interpreted.

4 Discussion

4.1 RQ1: psycho-emotional outcomes

No SAN subscale showed a statistically significant pre–post change in any group (all $p > .10$), indicating no detected within-lesson shift in psycho-emotional state associated with participation in the aromatherapy sessions. Because the SAN was administered at the beginning and end of the same lesson and averaged across sessions, this result concerns within-lesson change; it does not show whether multi-week exposure had a cumulative effect on psycho-emotional state, which would have required repeated assessments of baseline state across the study period. Because the design compared each participant's state before and after sessions rather than testing experimental against control classes on a shared outcome, this finding establishes an absence of detected pre–post change rather than a direct test of a treatment effect; a genuine effect could in principle have been masked by limited statistical power: with the achieved matched samples, the minimum detectable within-group effect ranged from d = 0.41 to d = 0.94 (§4.6). Given the confidence intervals, effects of up to about 0.7 standard deviations in either direction remain compatible with the data (Table 6; Figure 1), so the results indicate an absence of evidence for an effect rather than evidence of its absence.

This result contrasts with prior evidence that ambient essential oil exposure reduces anxiety and improves mood (Lehrner et al., 2000; Lehrner et al., 2005). However, those studies used brief, single-session designs in clinical or laboratory settings, quite different from the prolonged, repeated classroom exposure used here. If reduced odour perception developed over the course of

TABLE 8 Eysenck Personality Inventory (EPI) — Descriptive Statistics.

Group	*n*	E — Extraversion M (SD)	N — Neuroticism M (SD)	L ≥ 5 (%)	Interpretation
Schoolchildren — Zvyahel	13	10.85 (3.91)	15.85 (4.34)	31%	E: ambiversion; N: average
Schoolchildren — Zhytomyr	40	13.28 (3.71)	12.90 (5.33)	32%	E: ambiversion; N: average
All schoolchildren	53	12.68 (3.87)	13.62 (5.22)	32%	E: ambiversion; N: average
University students	11	10.36 (4.20)	17.00 (3.41)	36%	E: ambiversion; N: instability

E (Extraversion): 0–10 = introversion, 11–14 = ambiversion, 15–24 = extraversion. N (Neuroticism): 0–10 = stability, 11–16 = average, 17–22 = instability, 23–24 = borderline pathology. L ≥ 5 indicates potentially invalid responses (social desirability). EPI data reported descriptively only; no inferential statistics applied.

repeated sessions (§4.3), any psycho-emotional benefit tied to active olfactory stimulation would be expected to diminish accordingly — a possibility consistent with, though not proof of, the null result observed here. Moreover, because a 5-point rather than the original 7-point response format was used, absolute SAN scores are not directly comparable with normative data or with studies using the original format; our conclusions rest only on within-study comparisons.

4.2 RQ2: academic performance

Pre–post mathematics grade comparisons within experimental classes did not reveal consistent improvements attributable to the intervention; where grade improvements did occur, comparable or larger improvements were also observed in control classes that received no aromatherapy (§3.3). Because experimental and control classes were not randomly assigned, a genuine intervention effect would need to appear as differential improvement — experimental classes improving more than control classes — rather than as improvement common to both; the parallel pattern observed here does not provide evidence for a beneficial intervention effect; the primary Group × Time analysis is considered below.

The Group × Time interaction analysis (§3.4) provides a more direct test of differential improvement: no significant interaction was found at Zhytomyr, whereas at Zvyahel the control class improved more than the experimental classes. This result should not be read as evidence that aromatherapy reduced academic performance. An ANCOVA adjustment for baseline grade left the group effect essentially unchanged ($\beta = -0.33$ vs. unadjusted DiD = −0.37), indicating that regression to the mean is unlikely to account for the difference; design features provide alternative explanations that could not be separated from a treatment effect. Assignment to the control condition was not random: class 7B served as the control because several parents in this class did not consent to exposure, and, as a Grade 7 class, it covered different curricular content from the Grade 8 experimental classes during the study window, introducing a content-difficulty confound independent of aromatherapy exposure. Moreover, the effect was smaller, and would not survive multiplicity correction, when the Grade 5 class was excluded (Supplementary Table S2). Given these design features, the Zvyahel finding cannot be causally attributed to the intervention in either direction. Taken together with the paired-comparison results above, the data provide no evidence of a beneficial intervention effect on academic performance.

The absence of evidence for a beneficial intervention effect contrasts with prior studies reporting cognitive benefits of aromatherapy: Moss et al. (2017) found improved working memory in 9–11-year-old children following a 15-minute rosemary exposure in a classroom setting, and laboratory studies with adults similarly reported memory-enhancing effects of rosemary inhalation (Moss et al., 2003; Moss and Oliver, 2012). These studies differ from the present protocol in important respects: they used considerably shorter exposure durations (15–30 min vs. 25–80 min per session here) and fewer sessions. The brief single-session design used by Moss et al. (2017) may have been too short for olfactory habituation to develop, whereas the extended multi-session protocol used here created conditions under which repeated reductions in odour perception became evident (§4.3). Differences in exposure duration and session frequency may contribute to the divergent findings, although the oils, outcomes, and populations also differed between studies.

4.3 RQ3: intervention fidelity and reduced odour perception

Teachers and supervisors consistently reported that most participants ceased to perceive the scent within the first few minutes of each session (Table 1). Progressive dose escalation was employed to compensate, yet the scent repeatedly became imperceptible shortly after each increase; in the most extreme case, the dose increased from 7 to 33 drops (×4.7) in the university lavender phase without restoring stable perception. Subjective questionnaire data corroborated this pattern: half of the respondents (40/80) reported reduced scent perception during the first oil phase, a further 30% were uncertain, and 69% reported that the scent faded again following dose increases. Together, these observations indicate that perceived odour intensity could not be sustained despite active adjustment. This is a fidelity problem at the level of reported perception; airborne concentration and inhaled dose were not measured, so the data do not show whether physiological exposure declined.

Olfactory habituation, a well-documented consequence of repeated or prolonged odour exposure (Dalton, 2000; Pellegrino et al., 2017), offers one plausible account of this pattern, although the relative contribution of peripheral and central mechanisms remains debated (Mignot et al., 2022). Two features of the present protocol are consistent with this interpretation: rooms were ventilated before and after, but never during, sessions, so the within-session fading cannot be attributed to ventilation; and investigators informally noted that the scent was perceptible at the start of sessions before fading, although without concentration measurements this observation cannot establish that the dose was adequate. However, laboratory evidence also indicates that more intense and more trigeminal odorants tend to produce less habituation (Sinding et al., 2017), which would predict that dose escalation should at least partly restore perception; that it did not do so here is not fully explained by habituation alone. Moreover, initial doses were deliberately set near the perception threshold and below the room-diffusion dose suggested on the lavender and sweet orange oil labels, and airborne concentration was not measured. Because scent perception was assessed only through periodic questionnaires and teacher reports, central habituation cannot be separated from an insufficient airborne concentration relative to the perception threshold, and the habituation account should be regarded as an indirect interpretation. Finally, laboratory paradigms typically involve exposures lasting seconds to minutes, so their applicability to diffusion sustained over entire lessons is limited — a gap that the present classroom observations help to highlight.

Expectancy effects must also be considered when interpreting both the adverse reactions and the fidelity data. At Lyceum No. 28, complaints of headache or dizziness recurred among a small

number of pupils, while most participants reported not perceiving the scent at all (Table 1). Symptom reports in the apparent absence of a perceived odour are more readily explained by expectation than by a pharmacological effect: when the same chemical exposure is described as potentially harmful rather than benign, people report stronger odour and irritation and more health symptoms (Dalton, 1999; Dalton et al., 1997), and the relaxing effects attributed to lavender aroma have been shown to depend on participants' expectancies rather than on the aroma itself (Howard and Hughes, 2008). In the present study, pupils knew that an essential oil was being diffused, and classmates' complaints may have shaped expectations; the reported adverse reactions should therefore not be interpreted as evidence of a pharmacological effect of the oils at the doses used and may reflect, at least in part, nocebo responses. The same consideration applies to the fidelity argument: self-reported scent perception is itself susceptible to expectation and to the knowledge that diffusion was taking place, so both the reports of fading perception and the occasional reports of benefit are subjective indicators that cannot be validated without objective measures of exposure and odour perception.

4.4 Comparison with evidence from other educational settings

The null psycho-emotional findings contrast with evidence from other educational settings and student populations. In a randomised controlled trial with nursing students, aromatherapy, alone and combined with music, reduced perceived stress and improved self-esteem (Öner et al., 2026); lavender and rosemary oils reduced test-taking anxiety among graduate nursing students (McCaffrey et al., 2009); a meta-analysis of randomised trials concluded that inhalation aromatherapy alleviates test anxiety in college students (Luan et al., 2023); and a meta-analysis of lavender preparations reported anxiolytic effects, although most included trials were judged to be of low quality (Donelli et al., 2019). At the same time, the evidence is not uniformly positive: effects on cognitive performance have been weak, task-dependent, or absent in several studies (Ilmberger et al., 2001; Tong et al., 2025), consistent with the present results.

Several differences between these studies and the present one may account for the discrepancy. First, most positive studies used brief, attentive exposure, often immediately before or during a stressful event such as an examination, whereas the present study used passive ambient diffusion sustained over entire lessons — the condition under which reduced odour perception is most likely (Section 4.3). Second, positive studies typically measured state anxiety or perceived stress with instruments sensitive to acute change, whereas the SAN captures a broader functional state (well-being, activity, and mood). Third, randomised trials under controlled conditions reduce the variability that, in authentic classrooms, arises from lesson content, teacher behaviour, and — in the present case — wartime disruption. Fourth, participants in examination studies may expect benefit in a specific stressful situation, which can amplify self-reported effects. These differences suggest that the present null psycho-emotional findings concern within-lesson change under a specific delivery mode — prolonged passive diffusion — rather than aromatherapy as such.

4.5 Participant engagement and demand characteristics

Despite the absence of objective changes in psycho-emotional state or academic performance, teachers informally reported increased participant engagement during aromatherapy sessions, and self-reported cognitive benefits were occasionally noted (21% of participants) despite no corresponding change in academic outcomes. Plausible explanations include a Hawthorne effect (Roethlisberger and Dickson, 1939)—participants altering behaviour simply from being observed—and demand characteristics, particularly given that one questionnaire item was preceded by a statement of prior research findings (§3.6), which may have biased responses toward expected outcomes. These observations were not collected through standardised behavioural protocols and should be treated as corroborating field reports rather than systematic evidence.

4.6 Limitations

Several methodological limitations must be acknowledged. First, the quasi-experimental design with non-equivalent control groups precluded causal inference: classes were not randomised, and pre-existing differences between experimental and control groups cannot be ruled out, preventing valid attribution of post-intervention differences to the aromatherapy. Moreover, the intervention was delivered to whole classes, whereas inferential analyses treated pupils as independent observations; with only one control class per site, class-level clustering could not be modelled, so the reported p-values and confidence intervals may overstate the precision of between-group comparisons, and the Zvyahel grade finding in particular should be regarded as descriptive. Second, it was not possible to standardise oil concentration in the classroom air or to ensure equal exposure distances for all participants. Third, a protocol deviation occurred at Lyceum No. 28 where eucalyptus was replaced with sweet orange oil due to adverse reactions in five or more students in class 8D. Fourth, SAN responses could be linked across sessions only through the e-mail address recorded by the online form, and pre- and post-lesson averages were not always based on the same sessions, which reduced matched sample sizes. Fifth, participant absences due to wartime disruptions resulted in incomplete datasets and substantial attrition at some sites: matched SAN data were available for 24 of 88 enrolled experimental pupils at Lyceum No. 1, 46 of 55 and 25 of 29 at Lyceum No. 28, and all 11 university students. A sensitivity power analysis (two-tailed $\alpha = .05$, power = .80) indicated minimum detectable within-group effects of $d = 0.60$, 0.42, 0.58, and 0.94 for these samples, respectively, and $g = 0.71$ for the Zhytomyr SAN Group × Time comparison (46 vs. 25 participants); power to detect a small-to-moderate effect ($d = 0.30$) ranged from .15 to .51. The null results on the SAN subscales may therefore partly reflect insufficient statistical power rather than a true absence of effect. Sixth, the university

sample ($n = 11$) is too small for generalisation, and the absence of a control group at this site further limits interpretation. Seventh, the SAN questionnaire was administered using a 5-point response format rather than the original 7-point scale; although internal consistency in the present sample was high (Section 2.4.1), the validity of this modified version was not independently established, and its comparability with normative data based on the original format should be treated with caution. Eighth, one Activity-subscale item (item 16) was inadvertently omitted from the university sample's online questionnaire; a simulation in the school data indicated that pro-rating had only a minor effect on the Activity score (Section 2.4.1). Ninth, no independent compositional (e.g., GC–MS) analysis of the essential oils was performed. Finally, the study was conducted during an active period of armed conflict in Zhytomyr Oblast, Ukraine; chronic wartime stress is associated with impairments in attention, executive functioning, and emotional regulation (Arnsten, 2009) and with altered olfactory processing (Shin et al., 2023), and could plausibly have influenced both participants' psycho-emotional responses and the pattern of reduced odour perception documented in this study, though this could not be tested with the available data. Importantly, olfactory habituation was not directly measured in the present study and should therefore be regarded as an indirect interpretation based on repeated reductions in odour perception and progressive dose escalation. It should also be noted that the initial dose was deliberately calibrated near the threshold of perceptibility and was below the room-diffusion dose suggested on the lavender and sweet orange oil labels (approximately 3 drops per 10 m^2); investigators informally noted that the scent was perceptible at the start of sessions before fading. Because airborne concentration was not measured, this observation cannot establish that the starting dose was sufficient; an insufficient dose relative to the perception threshold therefore remains a plausible alternative or contributing explanation, and the relative contributions of dose calibration and habituation cannot be disentangled with the present data. In addition, control-group participation in the SAN assessment was inconsistent between sites: at Lyceum No. 28, the majority of control-class pupils completed the SAN questionnaire, whereas at Lyceum No. 1, only a small number of control-class pupils did so, likely reflecting a site-level misunderstanding that psycho-emotional assessment applied only to aromatherapy-exposed classes. This asymmetry further limits the validity of any between-site or between-group comparisons involving psycho-emotional state. Additionally, Zvyahel's control class (7B, grade 7) covered different curricular content than the grade-8 experimental classes during the study window, confounding class assignment with material difficulty independent of the intervention. Finally, a subset of the Zvyahel geometry grade records for class 8V (both pre- and post-intervention) were transcribed manually from a photographed paper grade book rather than sourced from a digital record as for all other classes; while cross-checked, this introduces a small, non-systematic risk of transcription error not present elsewhere in the dataset. Grade records were merged across subjects by pupil name; during revision, three duplicated records arising from spelling variants were identified and corrected, and all grade analyses were recomputed.

4.7 Directions for future research

The findings suggest several directions for future research. First, protocols should explicitly account for olfactory habituation: shorter sessions (10–15 min), active targeted inhalation rather than passive diffusion, or intermittent scent delivery may reduce the likelihood of olfactory adaptation and help maintain odour detectability throughout the intervention. Second, randomised controlled trials with larger samples, standardised air-concentration measurements, and individual-level data tracking are needed to permit causal inference. Third, the role of wartime stress and chronic psychological adversity as potential moderators of cognitive and emotional responses to aromatherapy warrants dedicated investigation, as wartime stress was not measured in the present study and may modify responses to aromatherapy.

4.8 Educational and practical implications

The findings caution against assuming that a sustained environmental stimulus remains psychologically available simply because it remains physically present in the classroom: perceived exposure diminished well before the intervention period ended, even as the physical stimulus continued. This distinction between physical presence and psychological availability carries several implications for educational practice. First, the lack of evidence for psycho-emotional benefits and of evidence for a beneficial academic effect under a passive, low-cost diffusion protocol suggests that schools should not adopt such interventions on the assumption of established efficacy; any classroom application of aromatherapy would benefit from evidence generated under conditions that address the fidelity limitations identified here. Second, this pattern has broader relevance for classroom-based sensory or environmental interventions generally: sustained, passive exposure—whether olfactory, auditory, or otherwise—may, although this was not tested here, be particularly vulnerable to adaptation effects that erode intended dosage over time, a consideration relevant to educators evaluating similar low-intensity, environment-based interventions.

5 Conclusion

This pilot study examined psycho-emotional and academic correlates of repeated passive aromatherapy in mathematics classrooms among 210 participants from three Ukrainian educational institutions. No SAN subscale showed a significant change from the beginning to the end of lessons (RQ1); because the SAN captured within-lesson change, these data do not show whether multi-week exposure had a cumulative psycho-emotional effect. For academic performance (RQ2), within-group pre–post comparisons showed no improvement attributable to the intervention beyond what was also observed in control classes; a Group × Time interaction analysis — the primary test of a differential intervention effect — likewise found no significant effect at Zhytomyr, while at Zvyahel the control class improved significantly more than the experimental classes (a result sensitive to analytic choices), a difference that cannot be causally attributed

to the intervention given non-random class assignment and curricular differences (§4.2). Despite repeated dose escalation, half of the questionnaire respondents reported reduced scent perception and a further 30% were uncertain, indicating that perceived odour intensity could not be maintained; this is consistent with olfactory habituation, although an insufficient airborne concentration cannot be excluded (RQ3). Sustained passive diffusion therefore appears difficult to maintain as a reliable classroom intervention, and future work should test shorter, active, or intermittent delivery protocols alongside objective measures of odour perception.

Data availability statement

The datasets presented in this article are not readily available because the dataset contains data from minor participants; access is subject to institutional ethical approval and data protection requirements. Requests to access the datasets should be directed to Alla Prus, prus@zu.edu.ua.

Ethics statement

The studies involving humans were approved by Academic Council of the Faculty of Physics and Mathematics, Zhytomyr Ivan Franko State University, Zhytomyr, Ukraine. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the adult participants (university students) and, for minor participants, by their legal guardians/next of kin.

Author contributions

AP: Writing – original draft, Data curation, Methodology, Conceptualization, Investigation, Project administration. OM: Validation, Writing – review & editing, Methodology, Resources. TK: Writing – review & editing, Validation, Formal Analysis. VS: Methodology, Writing – review & editing, Validation. OF: Writing – review & editing, Methodology.

Funding

The author(s) declared that financial support was not received for this work and/or its publication. The study was conducted within the departmental research project "Professional Competence of the Mathematics Teacher" (UkrINTEI registration No. 0123U100258), approved by the Academic Council of the Faculty of Physics and Mathematics, Zhytomyr Ivan Franko State University; this project did not provide dedicated funding.

Acknowledgments

The authors gratefully acknowledge Dr. Viktoriia I. Kravchenko, Ph.D. in Biological Sciences, Associate Professor at Taras Shevchenko National University of Kyiv, for her expert consultation on the selection of essential oils and the neurobiological rationale underlying this study. The authors also thank Maryna V. Malynovska, mathematics teacher, who conducted the intervention at Lyceum No. 1 named after Lesia Ukrainka (Zvyahel), and Alona A. Savytska, mathematics teacher, who conducted the intervention at Lyceum No. 28 named after Hetman Ivan Vyhovsky (Zhytomyr), for their contributions to intervention delivery and data collection.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. The authors used Claude (Anthropic; Claude Sonnet 5 and Claude Opus 5.5 models, accessed via Claude.ai) to assist with literature search and verification, re-checking of statistical computations from the raw data during revision, and language editing for clarity and fluency, given that the authors' first language is not English. All AI-assisted content was reviewed, verified, and edited by the authors, who take full responsibility for the accuracy and integrity of the manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2026.1982806/full#supplementary-material

References

Amland, T., Grande, G., Scherer, R., Lervåg, A., and Melby-Lervåg, M. (2025). Cognitive factors underlying mathematical skills: a systematic review and meta-analysis. *Psychol. Bull.* 151, 88–129. doi: 10.1037/bul0000457

Arnsten, A. F. T. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nat. Rev. Neurosci.* 10, 410–422. doi: 10.1038/nrn2648

Baron, R. A., and Thomley, J. (1994). A whiff of reality: positive affect as a potential mediator of the effects of pleasant fragrances on task performance and helping. *Environ. Behav.* 26, 766–784. doi: 10.1177/0013916594266003

Chang, C.-Y., Hsiao, M., and Chiang, W.-C. (2026). Mathematics performance predicts cognitive and affective math anxiety through mutual mediation pathways from adolescence onward with potential working memory moderations. *Sci. Rep.* 16, 10716. doi: 10.1038/s41598-026-45516-y

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2nd Edn. Hillsdale, NJ: Lawrence Erlbaum Associates.

Dalton, P. (1999). Cognitive influences on health symptoms from acute chemical exposure. *Health. Psychol.* 18, 579–590. doi: 10.1037/0278-6133.18.6.579

Dalton, P. (2000). Psychophysical and behavioral characteristics of olfactory adaptation. *Chem. Senses* 25, 487–492. doi: 10.1093/chemse/25.4.487

Dalton, P., Wysocki, C. J., Brody, M. J., and Lawley, H. J. (1997). The influence of cognitive bias on the perceived odor, irritation and health symptoms from chemical exposure. *Int. Arch. Occup. Environ. Health* 69, 407–417. doi: 10.1007/s004200050168

Danylchenko, N. O., and Vertel, A. V. (2012). *Psykhodiahnostyka: Praktykum [Psychodiagnostics: A Practicum]*. 2nd Edn. Sumy: Vyd-vo SumDPU im. A.S. Makarenka.

Donelli, D., Antonelli, M., Bellinazzi, C., Gensini, G. F., and Firenzuoli, F. (2019). Effects of lavender on anxiety: a systematic review and meta-analysis. *Phytomedicine.* 65, 153099. doi: 10.1016/j.phymed.2019.153099

Doskin, V. A., Lavrentieva, N. A., Sharai, V. B., and Miroshnikov, M. P. (1973). Test differentsirovannoi samootsenki funktsionalnogo sostoyaniya [test of differentiated self-assessment of functional state]. *Vopr. Psikhol.* 19, 141–145.

Eysenck, H. J., and Eysenck, S. B. G. (1964). *Manual of the Eysenck Personality Inventory*. London: University of London Press.

Eysenck, M. W., Derakshan, N., Santos, R., and Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion* 7, 336–353. doi: 10.1037/1528-3542.7.2.336

Howard, S., and Hughes, B. M. (2008). Expectancies, not aroma, explain impact of lavender aromatherapy on psychophysiological indices of relaxation in young healthy women. *Br. J. Health. Psychol.* 13, 603–617. doi: 10.1348/135910707X238734

Iglesias-Sarmiento, V., and Deaño, M. (2011). Cognitive processing and mathematical achievement: a study with schoolchildren between fourth and sixth grade of primary education. *J. Learn. Disabil.* 44, 570–583. doi: 10.1177/0022219411400749

Ilmberger, J., Heuberger, E., Mahrhofer, C., Dessovic, H., Kowarik, D., and Buchbauer, G. (2001). The influence of essential oils on human attention. I: alertness. *Chem. Senses* 26, 239–245. doi: 10.1093/chemse/26.3.239

Jellinek, J. S. (1997). Psychodynamic odor effects and their mechanisms. *Cosmetics Toiletries* 112, 61–72.

Johnson, A. J. (2011). Cognitive facilitation following intentional odor exposure. *Sensors* 11, 5469–5488. doi: 10.3390/s110505469

Kostka, J. K., and Bitzenhofer, S. H. (2022). How the sense of smell influences cognition throughout life. *Neuroforum* 28, 177–185. doi: 10.1515/nf-2022-0007

Lehrner, J., Eckersberger, C., Walla, P., Pötsch, G., and Deecke, L. (2000). Ambient odor of orange in a dental office reduces anxiety and improves mood in female patients. *Physiol. Behav.* 71, 83–86. doi: 10.1016/S0031-9384(00)00308-5

Lehrner, J., Marwinski, G., Lehr, S., Johren, P., and Deecke, L. (2005). Ambient odors of orange and lavender reduce anxiety and improve mood in a dental office. *Physiol. Behav.* 86, 92–95. doi: 10.1016/j.physbeh.2005.06.031

Luan, J., Yang, M., Zhao, Y., Zang, Y., Zhang, Z., and Chen, H. (2023). Aromatherapy with inhalation effectively alleviates the test anxiety of college students: a meta-analysis. *Front. Psychol.* 13, 1042553. doi: 10.3389/fpsyg.2022.1042553

McCaffrey, R., Thomas, D. J., and Kinzelman, A. O. (2009). The effects of lavender and rosemary essential oils on test-taking anxiety among graduate nursing students. *Holist. Nurs. Pract.* 23, 88–93. doi: 10.1097/HNP.0b013e3181a110aa

Mignot, C., Schunke, A., Sinding, C., and Hummel, T. (2022). Olfactory adaptation: recordings from the human olfactory epithelium. *Eur. Arch. Oto-Rhino-Laryngol.* 279, 3503–3510. doi: 10.1007/s00405-021-07170-0

Moss, M., Cook, J., Wesnes, K., and Duckett, P. (2003). Aromas of rosemary and lavender essential oils differentially affect cognition and mood in healthy adults. *Int. J. Neurosci.* 113, 15–38. doi: 10.1080/00207450390161903

Moss, M., Earl, V., Moss, L., and Heffernan, T. (2017). Any sense in classroom scents? Aroma of rosemary essential oil significantly improves cognition in young school children. *Adv. Chem. Eng. Sci.* 7, 450–463. doi: 10.4236/aces.2017.74032

Moss, M., Howarth, J., and Moss, H. (2025). Aroma of genius essential oil blend significantly enhances cognitive performance and brain metabolism in healthy adults. *Hum. Psychopharmacol.* 40, e70027. doi: 10.1002/hup.70027

Moss, M., and Oliver, L. (2012). Plasma 1,8-cineole correlates with cognitive performance following exposure to rosemary essential oil aroma. *Ther. Adv. Psychopharmacol.* 2, 103–113. doi: 10.1177/2045125312436573

OECD (2024). *Mathematics for Life and Work: A Comparative Perspective on Mathematics to Inform Upper Secondary Reform in England*. Paris: OECD Publishing. doi: 10.1787/26f18d39-en

Öner, U., Çevik Aktura, S., Özbay, H., and Erol, A. (2026). Effects of aromatherapy and music on stress and self-esteem in nursing students. *J. Nurs. Educ.* 65, 293–301. doi: 10.3928/01484834-20260129-02

Pellegrino, R., Sinding, C., De Wijk, R. A., and Hummel, T. (2017). Habituation and adaptation to odors in humans. *Physiol. Behav.* 177, 13–19. doi: 10.1016/j.physbeh.2017.04.006

Roethlisberger, F. J., and Dickson, W. J. (1939). *Management and the Worker*. Cambridge, MA: Harvard University Press.

Shin, M.-G., Bae, Y., Afzal, R., Kondoh, K., and Lee, E. J. (2023). Olfactory modulation of stress-response neural circuits. *Exp. Mol. Med.* 55, 1659–1671. doi: 10.1038/s12276-023-01051-x

Sinding, C., Valadier, F., Al-Hassani, V., Feron, G., Tromelin, A., Kontaris, I., et al. (2017). New determinants of olfactory habituation. *Sci. Rep.* 7, 41047. doi: 10.1038/srep41047

Spivak, L. M., and Osmanova, A. M. (2023). *Psykhodiahnostyka v Klinichnii Psykholohii: Navch. Posib. [Psychodiagnostics in Clinical Psychology: A Textbook]*. Kyiv: University "Ukraina".

Tong, Z., Liu, X., Song, Y., Guan, H., Wen, X., Hu, S., et al. (2025). The influence of odors on cognitive performance based on different olfactory preferences. *Build. Environ.* 271, 112603. doi: 10.1016/j.buildenv.2025.112603

Vance, D. E., Del Bene, V. A., Kamath, V., Frank, J. S., Billings, R., Cho, D.-Y., et al. (2024). Does olfactory training improve brain function and cognition? A systematic review. *Neuropsychol. Rev.* 34, 155–191. doi: 10.1007/s11065-022-09573-0

Wang, M., Mohd Matore, M. E. E., and Rosli, R. (2025). A systematic literature review on analytical thinking development in mathematics education: trends across time and countries. *Front. Psychol.* 16, 1523836. doi: 10.3389/fpsyg.2025.1523836