Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A school-based intervention to promote healthy lifestyles for cancer prevention in adolescents

Concetta Polizzi
University of Palermo
Flavia Caturano
Società Italiana di Psicologia Pediatrica (SIPPed)
Loredana Masseria
ASL Città di Torino
Alessandro Comandone
Medical Oncology Unit and Territorial Area of the Corporate Oncology Network
Carlo Picco
ASL Città di Torino
Giovanna Perricone
Ombudsman for Children and Adolescents of the Sicily Region
Flavia Altavilla
University of Palermo
Mariangela Sciandra
University of Palermo
Marco Speciale
marco.speciale02@unipa.it

University of Palermo

Article

Keywords: Adolescents, Cancer prevention, Health promotion, Topic modelling, Sentiment analysis

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11428850/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A school-based intervention to promote healthy lifestyles for cancer prevention in adolescents

Concetta Polizzi[1,2], Flavia Caturano[2,3], Loredana Masseria[3], Alessandro Comandone[4], Carlo Picco[3], Giovanna Perricone[5], Flavia Altavilla[1,2], Mariangela Sciandra[6], Marco Speciale[6*]

[1]Department of Psychology, Educational Science and Human Movement, University of Palermo, Viale delle Scienze, Building 15, Palermo, 90128, Italy.
[2]Società Italiana di Psicologia Pediatrica (SIPPed), Italy.
[3]ASL Città di Torino, Turin, Italy.
[4]Medical Oncology Unit and Territorial Area of the Corporate Oncology Network, Rete Oncologica del Piemonte e della Valle d'Aosta, ASL Città di Torino, Turin, Italy.
[5]Ombudsman for Children and Adolescents of the Sicily Region, Italy.
[6]Department of Economics, Business and Statistics, University of Palermo, Viale delle Scienze, Building 13, Palermo, 90128, Italy.

*Corresponding author(s). E-mail(s): marco.speciale02@unipa.it;

Abstract

Modifiable lifestyle factors such as smoking, alcohol consumption, poor diet and physical inactivity account for a large share of cancer risk, and these behaviours often take shape during adolescence. We evaluated "I Care for Me for You", a school-based cancer prevention programme in northern Italy grounded in Service Learning and Life Skills education, in which students attended expert-led sessions and then designed health promotion campaigns for peers and families. In a quasi-experimental design with unlinked pre-intervention (n = 1,225) and post-intervention (n = 384) samples of students aged 11–19 years, representations of health were elicited with a narrative instrument and the open-ended responses were analysed through keyword coding, lexicon-based sentiment analysis and Latent Dirichlet Allocation with bootstrap confidence intervals. Baseline knowledge was high, yet only weakly associated with self-reported healthy behaviour. After the intervention, responses on the role of the environment shifted from delegation to school and family towards awareness-raising and education, and

responses on personal action shifted away from generic references to healthy lifestyles. Overall, 82% of students rated the programme as useful and 63% reported intending to change an unhealthy behaviour. Involving adolescents as health promoters may help bridge the gap between health knowledge and action.

Keywords: Adolescents, Cancer prevention, Health promotion, Topic modelling, Sentiment analysis

1 Introduction

Cancer represents one of the leading causes of morbidity and mortality worldwide. According to estimates from the World Health Organization, approximately 20 million new cases of malignant neoplasm and 9.7 million related deaths were recorded in 2022 [1].

The Italian Association of Medical Oncology (AIOM) reported 390,100 new diagnoses, with breast, colorectal, lung and prostate cancers representing the most frequent malignancies in the Italian context [2]. Projections from the American Cancer Society anticipate 35 million new cases globally by 2050, driven primarily by population aging [3]; however, the World Cancer Research Fund (WCRF) emphasizes that modifiable risk factors, including tobacco use, alcohol consumption, obesity, physical inactivity and poor diet, also contribute substantially to cancer pathogenesis through genetic and inflammatory pathways [4]. Notably, adherence to the evidence-based prevention recommendations of the WCRF and the American Institute for Cancer Research (WCRF/AICR) has been associated with a 10–61% reduction in cancer incidence and mortality [4, 5] and the 2021 European Cancer Plan estimates that 40% of cancers are preventable through healthy lifestyle promotion and early screening [6].

Adolescence constitutes a critical "window of susceptibility" during which risk behaviors, tobacco use, physical inactivity, unbalanced diet, alcohol misuse, unprotected sexual activity and excessive UV exposure, become established and persist into adulthood [7, 8]. Despite this vulnerability, adolescents demonstrate notable gaps in cancer-risk literacy: they rarely identify obesity or low-fiber diets as risk factors and while awareness of smoking-related harms is relatively high, understanding of the protective roles of diet, physical activity and vaccination remains limited [9–11].

A recurring finding across literature is the discrepancy between theoretical health knowledge and actual behavior. Adolescents often associate health with body aesthetics and weight conformity rather than long-term disease prevention and immediate gratification frequently overrides awareness of health risks [12, 13]. This gap is compounded by peer pressure toward rigid aesthetic ideals, pervasive exposure to misleading media and influencer content and online misinformation that undermines public health messaging [14, 15].

In addition, previous studies have consistently reported that substantial emotional barriers impede the adoption of protective behaviors and the seeking of medical assistance [16, 17]. Among these barriers, embarrassment related to physical exposure along with apprehension about potential medical findings and concern over alarming

parents, leads many adolescents to disregard unusual symptoms, favoring a "wait and see" approach.

Beyond social and emotional determinants, cognitive functioning plays an important role in shaping health-related representations. Executive functions, responsible for planning, monitoring and goal-directed behavior, are linked to longer life expectancy and reduced chronic disease incidence [18] and are associated with an analytical thinking style [19]. Weak executive functioning and impulsivity, conversely, correlate with earlier substance use, unhealthy eating and obesity [20].

The present study conceptualizes four cognitive styles relevant to health representations: the analytical style, which supports risk evaluation and long-term planning but may produce rigid thinking; the concrete/experiential style, which facilitates practical behavior change through direct experience but struggles with delayed or non-observable risks; the abstract style, which supports comprehension of long-term prevention through logical reasoning but risks theoretical detachment and the inventive style, which fosters flexible, creative health behaviors but may lack the consistency required for sustainable habits.

Building on this evidence base, the school-based intervention program "I Care for Me for You" was implemented with 4,746 lower and upper secondary school students in the Piedmont and Aosta Valley regions of Italy. Distinctively, the program integrates primary prevention (healthy lifestyle promotion and human papillomavirus (HPV) vaccination) with secondary prevention (cancer screening promotion), employing a co-constructed educational approach that positions adolescents as active agents rather than passive recipients of health information.

The present study aims to evaluate the intervention program underlying the project "I Care for Me for You" within the Italian school context, with regard to changes in adolescents' representations of health and healthy lifestyles. In order to promote the adoption of healthy lifestyles, it is essential to assess and intervene on the cognitive beliefs of adolescents regarding such behaviors [21].

Representations of health and healthy lifestyles grounded in realistic, scientific and applicable data are assumed to promote adolescent agency, increasing confidence in personal capabilities and fostering a conscious, intentional and responsible attitude toward health. Furthermore, considering that such potential transformative effects of the program may be influenced by socio-demographic variables such as age, sex and nationality, as well as by the predominant cognitive style in approaching problematic situations, the present study developed the following hypotheses:

H1: It is hypothesized that the project "I Care for Me for You" promoted, among participants, the development of representations concerning health management and healthy behaviors oriented toward agency. Specifically, it is hypothesized that representations centered on individual responsibility and critical thinking in evaluating approaches to health developed. Furthermore, such representations may vary as a function of the variables age, sex, and nationality.

H2: It is hypothesized that a possible correlation exists between representations of health and healthy behaviors and the cognitive style characterizing problem-solving approaches.

2 Methods

The study adopted a quasi-experimental pre-post research design to examine changes in representations of health and healthy behaviors associated with the intervention program developed within the project "I Care for Me for You".

2.1 Participants

The project was conducted during the 2024/2025 school year, involving 52 secondary schools in the Piedmont and Aosta Valley regions, of which 23 were lower secondary schools and 29 were upper secondary schools. The project was promoted by the Local Health Authority of the City of Turin (ASL Città di Torino), with the involvement of a multiprofessional team of experts from the Piedmont and Aosta Valley Oncology Network and with the support of the CRT Foundation.

No exclusion criteria were established for students; consequently, the project was offered to all classes within the participating schools.

The total number of students who participated in the project "I Care for Me for You," including partial participation, was 4,746. However, for the purposes of the present study, only data from students who completed the baseline assessment, conducted a few days prior to the start of the intervention program, were considered.

The initial pre-intervention dataset comprised 1,234 students; after excluding cases in which the baseline instruments were not fully completed or demographic information was missing or inconsistent, the analysable pre-intervention sample comprised 1,225 students, the vast majority of Italian nationality. Specifically, the sample included 444 lower secondary school students, aged eleven to thirteen years, and 781 upper secondary school students, aged fourteen to nineteen years (Table 1).

At the post-intervention stage, upon completion of the program, only 384 students correctly and fully completed the assessment instruments, the vast majority of whom were of Italian nationality. This sample comprised 146 lower secondary school students, aged eleven to thirteen years, and 238 upper secondary school students, aged fourteen to nineteen years. Compared with the pre-intervention sample, the post-intervention sample included a higher proportion of males (62.5% vs 48.1%), whereas the distributions by school level and nationality were similar (Table 1).

Table 1 Characteristics of the pre-intervention and post-intervention samples.

		Pre (N = 1225)		Post (N = 384)	
Participants		**n**	**%**	**n**	**%**
Age	11–13 years	444	36.2	146	38.0
	14–19 years	781	63.8	238	62.0
Sex	Male	589	48.1	240	62.5
	Female	636	51.9	144	37.5
Nationality	Italian	1087	88.7	346	90.1
	Foreign	138	11.3	38	9.9

2.2 Ethics declarations

The study was reviewed by the Ethics Committee of the Italian Society of Pediatric Psychology (Società Italiana di Psicologia Pediatrica, SIPPed), which expressed a favourable opinion (prot. no. 20, 22 December 2025). The opinion was issued after the completion of data collection, which took place during the 2024/25 school year within the monitoring and evaluation plan of the programme, adopted by the participating schools as part of their institutional educational plan. For participants under 18 years of age, informed consent was obtained from parents or legal guardians, together with the assent of the students; participants aged 18 or older provided their own informed consent. Participation was voluntary and the questionnaires, accessed via QR code, were completed anonymously, without personal identifiers. All procedures were performed in accordance with the Declaration of Helsinki and relevant national guidelines and regulations.

2.3 Measures and Procedures

The study was structured into three sequential phases:

1. *Baseline* (T0): in the days immediately preceding the start of the project activities, several instruments were administered to define the initial profile of participants with regard to pre-existing knowledge of health and healthy lifestyles, adherence to key recommendations on healthy lifestyles and predominant cognitive styles in problem-solving management. In addition, a narrative instrument was administered to assess the baseline representations of health and the management of healthy lifestyles held by participants.
2. *Intervention*: implementation of the cancer prevention pathway, described in detail in Section 2.3.2.
3. *Post-Test Assessment* (T1): the narrative instrument used to assess representations of health and the management of healthy lifestyles was administered again upon completion of the intervention, in order to evaluate potential effects of change.

2.3.1 Baseline Phase (T0)

In order to define the initial profile of participants with regard to pre-existing knowledge of health and healthy lifestyles, adherence to key recommendations on healthy lifestyles and predominant cognitive styles in problem-solving management, the following instruments were administered:

- A questionnaire titled "*What I know about health...*", developed by the research team, designed to assess the knowledge of cancer prevention and healthy lifestyles held by participants. Comprising twelve multiple-choice items (each with a single correct response consistent with World Health Organization recommendations), the instrument investigated knowledge regarding the definition of health, the influence of social media on food and substance use, the cancer risk associated with smoking, physical inactivity, unhealthy diet and the utility of vaccination among other topics.
- A questionnaire titled "*Healthy Lifestyles*", developed in accordance with World Health Organization recommendations on diet, physical activity, technology use and

beverage consumption (for example: "How many servings of fruit do you usually eat per day? How many minutes per day do you spend engaging in physical activity? How many times per day do you drink fruit juices or sugar-sweetened beverages?"). This instrument enabled the assessment of the adherence of participants to key recommendations on healthy lifestyles.

- The P.I.A. ("Prove di Intelligenze Applicate") [22], a validated and standardized instrument designed to assess the cognitive styles predominantly employed by participants of both age groups in problem-solving contexts. As previously noted, cognitive styles in problem-solving management may influence healthy lifestyle behaviors and representations of health. Accordingly, the cognitive style profile of the sample was assessed in order to examine potential correlations with representations concerning the management of healthy lifestyles and to test the second research hypothesis. The P.I.A. is designed to assess the preferential mental processes employed by an individual in addressing and managing specific problem-solving situations. The processes investigated are identified as forms of "applied intelligence," which contribute to individual differences in cognitive style, defining distinct modes of intelligent functioning. In its validated form, the P.I.A. is structured as a simulation activating "safety systems" [23], whereby participants engage with a narrative in which a character, Roberto, is the protagonist of a journey and the author of the choices made, encountering difficulties along the way. The instrument comprises a sequence of twelve problem-solving situations (twelve items), each corresponding to a specific stage of the journey undertaken by Roberto, with the aim of discovering a treasure that is ultimately revealed to correspond to the identification of the preferred processing tendency of the respondent (analytical, inventive, concrete or abstract) in managing problem-solving situations. For each of the twelve problem-solving situations constituting the items of the instrument, respondents are required to select one of four response options, each individually identified by the reference model as an indicator of analytical, concrete, inventive or abstract intelligence.

To test the first research hypothesis, concerning changes in representations regarding the management and promotion of healthy lifestyles, the baseline assessment included the administration of a narrative instrument denominated "*An INFO to express my view on healthy lifestyles related to cancer prevention*" (hereafter INFO), developed by the research team (unpublished). The instrument was administered via a Google Forms questionnaire, completed anonymously by students on their own smartphones by accessing a dedicated QR code.

This narrative instrument enabled the exploration, at the pre-intervention stage (with re-administration at the post-intervention stage), of the representations regarding health and prevention held by students, as well as their emerging needs. Specifically, the instrument presented students with twelve items of health-related information concerning health and healthy lifestyles related to cancer prevention, developed and provided by nationally and internationally accredited scientific institutions in the healthcare and educational fields. The 12 items (Info units) concerned, in order: healthy lifestyles and cancer prevention (Info 1), marketing (Info 2), smoking (Info 3), diet and body weight (Info 4), physical activity (Info 5), processed foods

(Info 6), red meat (Info 7), sugary and alcoholic beverages (Info 8), dietary supplements (Info 9), sun protection (Info 10), HPV vaccination (Info 11) and online health information (Info 12). For each item of information presented, students were asked to express their own views, specifying what they believed peers of the same age thought about the information provided, what role the environment (family, school, healthcare services, sports centers and so forth) should play in facilitating the adoption of healthy lifestyles among adolescents and what indications or suggestions they themselves derived from the information provided. Each item was accordingly followed by three specific questions: "What do peers your age think about this information?"; "What should the environment do to facilitate its application among adolescents?" and "What should adolescents change about their lifestyle in light of the information provided?". These characteristics render the INFO instrument capable of activating adolescent agency, understood as the capacity to make autonomous choices and decisions regarding important life issues, such as health.

Within this school-based prevention pathway, the INFO instrument therefore represented not only a fundamental assessment tool for evaluating knowledge, competencies and attitudes essential to healthy lifestyles, but also an intervention tool in itself, insofar as the reflective questions accompanying each item enabled respondents to activate specific cognitive, emotional, and relational skills. For instance, being prompted to consider the perspectives of same-age peers regarding a given health-related issue may promote important skills such as critical thinking, perspective taking and creative thinking. In this sense, the INFO instrument constitutes, in itself, a potential tool for change among the adolescents involved in the pathway. In particular, the instrument facilitates the activation of life skills and soft skills, whereby the former refer to a set of abilities acquired through life experiences or education that foster a positive and flexible attitude, enabling individuals to manage the difficulties of life [24], while the latter refer to intrapersonal and interpersonal competencies essential for personal development, social participation, and life success, influencing the manner in which individuals live and interact with others [25].

As an example, Info 1 read as follows: *"According to the Italian Ministry of Health, up to 40% of cancers can be prevented by adopting correct lifestyles and adhering to the cancer screening programs offered by the National Health Service. Many cancers are potentially preventable, as they are associated with modifiable risk factors, including tobacco use, unhealthy diet, harmful and risky alcohol consumption, insufficient physical activity and sedentary behavior. Adopting a healthy and active lifestyle represents the first step toward reducing cancer risk."*

2.3.2 Intervention Phase

The intervention program implemented within the project "I Care for Me for You" was designed to promote, among adolescents, representations functional to actively and adaptively managing the relationship with the environment, orienting them toward interpretations and perspectives on lifestyles that support the selection and adoption of healthy behaviors.

From a methodological standpoint, the intervention is grounded in the pedagogical model of "Service Learning" [26, 27], which integrates active, practical learning with

a service provided to the community. Indeed, a fundamental aspect of adolescent development within the program was the assignment of a task of high social value: the design and implementation of a health promotion campaign directed at peers and families, taking into account the information provided by experts and teachers, as well as the knowledge acquired or reinforced regarding healthy lifestyles functional to reducing potential cancer risk.

In addition, the scientific methodology of "Life Skills" [24] was employed, fostering critical thinking with respect to persuasive communication strategies used by tobacco, alcohol and food industry corporations, as well as in the analysis of health-related fake news. Further competencies promoted included decision making, understood as the capacity to make informed choices regarding healthy lifestyles and effective communication, with particular reference to assertiveness in the context of peer pressure [24].

Through the integration of these approaches and methodologies, students are enabled to become health promoters themselves.

From a procedural standpoint, the project comprised an informational component delivered by a multiprofessional team (including oncologists, nurses, psychologists, nutritionists and health journalists) and by teachers. Given that this information was intended to promote the reinforcement or development of representations capable of exerting a positive impact on the selection of healthy lifestyles, information provided was realistic, accessible and inclusive [6, 28]. The informational component was followed by a practical component in which students designed and implemented health promotion campaigns for the benefit of the community.

More specifically, the intervention program comprised two interactive sessions, each lasting two hours, involving a multiprofessional team of experts from the Piedmont and Aosta Valley Oncology Network (including oncologists, nurses, psychologists, nutritionists and health journalists), together with teachers from the participating schools. These sessions with experts were interspersed with a "homework" phase conducted in the classroom under teacher supervision and were followed by community events.

The first session, of an informational nature, was designed to provide clear guidance on cancer prevention and the promotion of healthy lifestyles, supported by audiovisual materials. Specifically, the topics addressed included the National Health Service (its structure and the provisions in place to protect the health of citizens); persuasive communication strategies adopted by tobacco, alcohol and food industry corporations from the 1950s to the present; the process of consumer loyalty formation beginning in childhood; the influence of peers and family on health-related behaviors; oncological pathologies; cancer prevention through screening programs offered by the National Health Service; the value of vaccination, with particular reference to human papillomavirus vaccination and fake news concerning health-related topics.

During these presentations, students were encouraged to interact actively with cancer prevention experts through questions and reflective discussion. Subsequently, over a period of approximately two months, students at school, coordinated and supervised by teachers, examined the informational material received in greater depth and conducted further research on healthy lifestyles, drawing on additional resources, including multimedia materials. On this basis, students proceeded to conceive and

develop awareness and health promotion campaigns directed at both peers and families. Students were able to select among various creative formats for the development of these promotional campaigns, including posters, songs, videos and thematic stamps.

The second session involved the presentation of the health promotion campaigns developed by students to the team of experts and teachers, accompanied by further discussion and elaboration of the topics addressed. Following this extensive classroom discussion, the health promotion campaigns developed by students were refined and subsequently disseminated to the community, for example through the websites of local health authorities. The pathway concluded with the organization of community events held in the main public squares of the cities in which the participating schools were located, with the aim of presenting the project and providing visibility to the work of students to the general public and to political governance bodies.

2.3.3 Post-Test Assessment Phase (T1)

Approximately ten days after the conclusion of the intervention, the narrative instrument "INFO", described above, was administered again, in order to conduct a pre and post-intervention comparison of representations concerning health and the management of healthy lifestyles.

A satisfaction questionnaire, denominated "Human Satisfaction" and developed for the present project, was also administered. This questionnaire comprised three items designed to assess, among adolescents: satisfaction with participation in the project; the predominant emotional state experienced throughout the intervention pathway, with response options including pleasure, sadness, disgust, indifference or boredom, and anger; and the intention, following the intervention, to modify a personal behavior identified, as a result of the project, as unhealthy.

2.4 Data Analysis

The analysis followed the sequence of instruments. Section 2.4.1 describes the analysis of the baseline questionnaires and the coding of the skills and needs activated in the INFO responses; Section 2.4.2 describes the sentiment and content-category analyses; Section 2.4.3 describes the topic model used to detect pre-post thematic shifts; and Section 2.4.4 describes the analysis of cognitive style data collected through the P.I.A. instrument.

The INFO instrument comprises 12 Info units on health conditions and lifestyle habits, each followed by the same three open-ended questions, yielding 36 items per respondent. Responses to the three questions formed three separate corpora: D1, addressing peer perceptions of each Info unit; D2, addressing the role expected of the environment in supporting the suggested lifestyle; and D3, addressing the actions respondents themselves might take. Given the open-ended format and the size of the corpora, a computational approach to text analysis was adopted, and pre- and post-intervention texts were analyzed jointly within each corpus to preserve a shared semantic space.

Because the instrument was administered anonymously, no identifier links a student's pre- and post-intervention responses. The two waves were therefore treated

as independent samples drawn from the same population at different points in time, although some students may have contributed to both, and all pre-post comparisons were conducted at the aggregate level: they describe differences between two groups and do not support inferences about change within individual students.

Prior to analysis, missing responses and non-informative content, that is, answers lacking substantive content such as "I don't know," generic expressions, and isolated monosyllables, were removed. As the INFO questionnaire progressed, a fatigue effect emerged, reflected in a non-response rate that rose in later sections (Fig. 1a) and in responses that grew shorter (Fig. 1b); the quantity and quality of available information therefore varied across items, and this pattern was taken into account when interpreting the results. Response length also showed a marked positive skew, with mostly short responses and comparatively few extensive ones, a typical feature of open-ended textual data that justified the use of statistical procedures robust to non-Gaussian distributions.

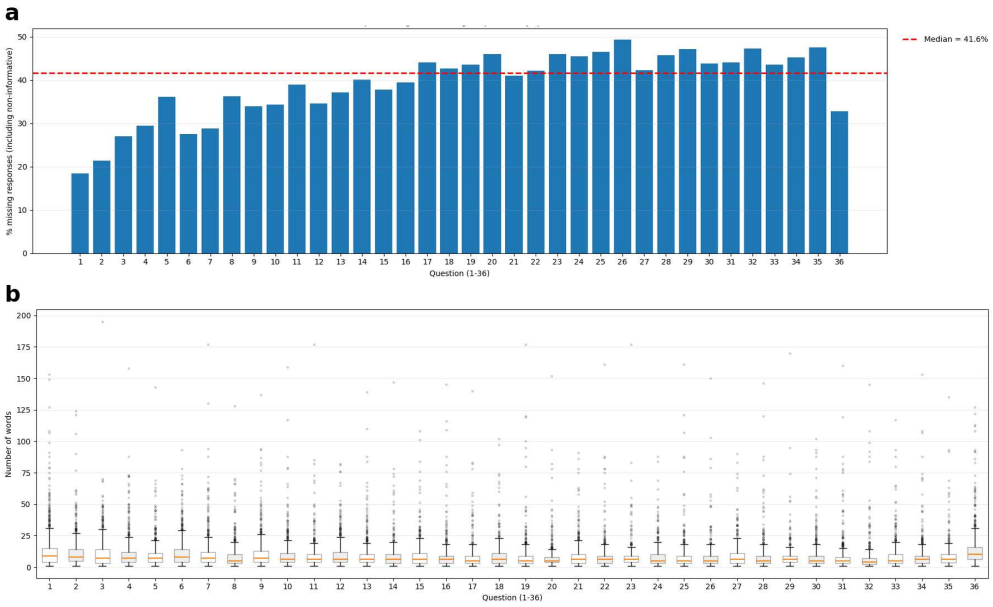

Fig. 1 Response quality across the 36 items of the narrative instrument. a, Percentage of missing or non-informative responses for each item; the dashed line indicates the median (41.6%). **b**, Distribution of response length, in number of words, for each item, excluding missing and non-informative responses.

Participating students were also nested within classes and schools (52 schools in total), a hierarchical structure that the present analyses do not model explicitly. Respondents are treated as independent units, although students sharing a class or school may share sources of variability such as the teacher, the local socioeconomic context, or the mode of administration. In the absence of random-effects models or

cluster-robust standard errors, the reported confidence intervals and P values may underestimate the true uncertainty of the estimates.

2.4.1 Baseline questionnaires and coding of skills and needs

Baseline knowledge of health and healthy lifestyles was assessed on the pre-intervention sample (N = 1,225) through the "What I know about health..." questionnaire, which comprises 12 multiple-choice items with a single correct answer each. The total score, obtained by summing correct responses, ranges from 0 to 12 and was classified into three bands: low (0 to 3), medium (4 to 7), and high (8 to 12). Scores were compared by sex, school level, and nationality using analysis of variance.

Adherence to healthy lifestyles was assessed through the Healthy Lifestyles Questionnaire, comprising 7 items on diet, physical activity, technology use and beverage consumption, with a theoretical range of 9 to 18 and higher scores indicating greater adherence. Scores were compared by sex, age group, and nationality using analysis of variance, and the association between knowledge and adherence was estimated through Spearman's rank correlation. The significance level was set at $\alpha = 0.05$, and correlation tests were two-sided.

Valid responses to the three questions following each of the 12 Info units were analyzed for the skills and needs they activated. Responses to D1 and D2 were examined for the cognitive and relational skills involved, drawing on the WHO life skills framework [24] and on the soft skills literature [25]; responses to D3 were examined for the needs students expressed regarding healthy lifestyles, since this question most directly elicits their agency and capacity to voice an opinion. For D1, seven skills were considered: six of the ten WHO life skills, namely critical thinking, relationship management, emotional management, problem solving, decision making, and creative thinking, together with planning. For D2, three soft skills were considered, namely perspective taking, evaluating, and classifying. For D3, expressed needs were coded into nine categories: information, knowledge, experience, listening, support, recognition, typicality, atypicality, and scaffolding. All categories were identified through a keyword-based system derived from this theoretical framework, with relative frequencies calculated conditional on the activation of each skill or need, and their distribution was examined by sex and school level, comparing lower and upper secondary students.

2.4.2 Sentiment and content categories

Sentiment analysis used natural language processing to identify and quantify the emotional content of students' responses to each Info unit. It drew on Plutchik's model of emotion [29, 30], which organizes emotional experience into eight primary emotions: joy, trust, fear, surprise, sadness, disgust, anger, and anticipation. Emotions were assigned to the texts using the Italian-adapted NRC Emotion Lexicon (EmoLex) [31], which links each word to one or more emotion categories. The analysis was run separately for D1, D2, and D3, applying the same classification procedure throughout so that the resulting emotional profiles would be comparable across corpora.

For each Info unit, the emotion associations of the words occurring in the responses were counted and aggregated by polarity, with joy and trust classified as positive and sadness, disgust, anger, and fear as negative; the positive and negative shares

were expressed as percentages of all positive and negative associations. Surprise and anticipation were assigned to the positive pole for consistency with Plutchik's original structure; since both are better characterized as activating than as intrinsically positive, this assignment should be regarded as a simplification.

Responses to all three questions were also coded into thematic content categories derived inductively from the corpus. Category distributions were examined for six Info units selected as reference cases on the basis of their content, covering the main risk domains addressed by the intervention: marketing (Info 2), processed foods (Info 6), HPV vaccination (Info 11), diet and body weight (Info 4), sugary and alcoholic beverages (Info 8) and dietary supplements (Info 9). Percentage distributions were compared by sex and school level.

2.4.3 Topic modeling (Latent Dirichlet Allocation)

To detect systematic thematic shifts between pre- and post-intervention responses, and thereby examine differences associated with the intervention, Latent Dirichlet Allocation (LDA) [32] was applied. LDA is a probabilistic generative model that represents each document as a mixture of latent topics and each topic as a distribution over the vocabulary. The model was estimated on the joint corpus of pre- and post-intervention responses from lower and upper secondary students, yielding a single shared thematic space. The number of topics was fixed at $K = 10$, chosen for substantive interpretability rather than through a quantitative optimization criterion; for each of the three questions, this value yielded a set of distinguishable and interpretable topics, consistent with the representative keywords reported in Tables 2–4.

In line with the independent-samples design described above, comparisons were conducted at the aggregate level, with the document as the unit of analysis, so that variations in mean topic proportions reflect shifts in the overall distribution of the sample rather than change within the same individuals. A thematic shift was considered robust only when the 95% bootstrap confidence interval excluded zero in both the unweighted and the length-weighted means, reducing the risk that observed differences reflect the distribution of text length rather than a genuine reallocation of thematic content between conditions. No correction for multiple comparisons was applied across the 30 topic-level comparisons (ten topics for each of the three questions); the double criterion was adopted to limit spurious findings, and these comparisons are interpreted as exploratory.

2.4.4 Analysis of cognitive style data

To explore hypothesis H2, and examine the relationship between cognitive style in problem solving and the representations expressed in students' texts, INFO data were integrated with the P.I.A. instrument (Prove di Intelligenze Applicate), administered across school levels. For both subsamples, students' cognitive styles were classified into the instrument's four intelligences, analytical, concrete, inventive, and abstract, and the distribution of predominant styles was described separately for the two school levels.

To explore whether cognitive functioning was associated, at the aggregate level, with INFO textual content, an additional exploratory bootstrap procedure was used.

As no individual identifiers link psychometric measures to a given student's INFO responses, this is not an individual-level correlation, and its results cannot be read as evidence for or against a genuine association; the procedure assumes the two samples are drawn from the same school population, a plausible but empirically unverifiable assumption. Correlations were estimated separately for LDA topic proportions and sentiment indicators, using bootstrap resampling with replacement on both datasets. Results are therefore presented as exploratory only, not as a formal test of H2.

3 Results

We first describe participants' baseline knowledge of and adherence to healthy lifestyles. We then report the representations of health elicited through the INFO narrative instrument, namely the skills and needs activated, the sentiment expressed, the content categories identified and the thematic differences between the pre- and post-intervention samples. Finally, we examine the exploratory relationship between cognitive style and these representations, and students' satisfaction with the programme.

3.1 Sample profile in terms of knowledge about and adherence to healthy lifestyles

Students' baseline knowledge, assessed via the "What I know about health..." questionnaire, showed a mean score of 9.60 (SD = 2.13, Mdn = 10.0), placing the sample in the high band, where the distribution was in fact concentrated, with 84.9% of respondents scoring high (8 to 12), 13.14% medium (4 to 7), and 1.96% low (0 to 3). Among sociodemographic variables, analysis of variance showed significant differences only by sex ($F = 27.28$, $P < 0.001$) and nationality ($F = 28.52$, $P < 0.001$). Females outscored males in both age groups (11 to 13 years, M = 9.8, SD = 1.8 vs M = 9.1, SD = 2.5; 14 to 19 years, M = 10.1, SD = 1.8 vs M = 9.2, SD = 2.3), and Italian students scored higher than foreign students (M = 9.7, SD = 2.0 vs M = 8.7, SD = 2.7).

Adherence to healthy lifestyles, assessed at baseline, showed a mean score of 13.73 (SD = 3.58, Mdn = 14.0, N = 1,225), within the intermediate range of the 9 to 18 scale. Analysis of variance revealed significant differences by age ($F = 4.26$, $P = 0.03$) and sex ($F = 4.14$, $P = 0.04$). Lower secondary students, aged 11 to 13, scored higher than upper secondary students, aged 14 to 19, with respective means of M = 14.37, SD = 3.5 and M = 13.3, SD = 3.6. Males scored higher than females in both age groups, more markedly among older students (11 to 13 years, M = 14.4, SD = 3.5 vs M = 14.3, SD = 3.5; 14 to 19 years, M = 13.7, SD = 3.4 vs M = 12.9, SD = 3.6). Nationality showed no detectable effect, though foreign male students tended to score higher than Italian males across both age groups, while among females Italian students scored higher at the younger age and foreign students at the older age.

Finally, lifestyle adherence and knowledge scores correlated weakly overall (Spearman's $\rho = 0.067$, $P = 0.019$, n = 1,225). This association was present among females ($\rho = 0.157$, $P = 7 \times 10^{-5}$, n = 636), lower secondary students ($\rho = 0.169$, $P = 3.5 \times 10^{-4}$, n = 444), and foreign students ($\rho = 0.170$, $P = 0.046$, n = 138), but absent among males and Italian upper secondary students, a pattern suggesting the

knowledge-behaviour link is strongest in subgroups with less prior exposure to health information. Overall, participants began the project with a solid baseline of health knowledge and a reasonably good adherence to healthy behaviours, as reflected in their own reports of diet, physical activity, and related habits.

3.2 Representations concerning health management and healthy lifestyles

The following analyses concern the responses to the INFO instrument for the three questions (D1, D2 and D3) described in the Methods.

3.2.1 Cognitive and relational skills activated, and needs emerging through the INFO instrument

For D1, the most frequently emerging life skills are critical thinking, relationship management, and problem solving; decision making and planning are less represented, and creative thinking and emotional management carry only marginal weight. Across nearly all skills, active documents show a predominantly male composition, generally between 55% and 60%, the one exception being creative thinking, where females show a slight prevalence (Fig. 2a). Upper secondary students predominate throughout, with shares ranging from 58% to 75% depending on the skill (Fig. 2b).

For D2, perspective taking shows the highest activation rate, followed by classifying and evaluating. The same male prevalence observed for D1 recurs here, at 62.65% for perspective taking and 61.28% for classifying (Fig. 3a), alongside a marked predominance of upper secondary students, at 69.67% and 64.63% respectively (Fig. 3b).

For D3, most need categories show a male overrepresentation, reaching 67.29% for information, 78.69% for knowledge, and 64.24% for listening. Experience runs counter to this trend, with a slight female majority (52.27%). Recognition and atypicality, both very low in absolute frequency, consist entirely of male contributions, while support, the most frequent need overall, shows a more balanced but still male-leaning distribution (Fig. 4a). Upper secondary students again predominate, from 72.12% for information to 84.09% for experience (Fig. 4b).

3.2.2 Sentiment profiles

Sentiment activated through the INFO instrument was analysed separately for each of the three questions attached to every Info unit. For each Info unit, the shares of positive and negative emotion associations were computed as described in the Methods.

For D1, positive sentiment predominates across most Info units, with two clear exceptions, Info 7 (64.46% negative) and Info 3 (64.36% negative), while Info 8, Info 4 and Info 6 sit close to equilibrium. The strongest positive values appear for Info 10 (84.46%), Info 5 (72.50%) and Info 2 (65.49%) (Fig. 5a).

D2 skews even more strongly positive, with less internal dispersion than D1. Info 10 (85.79%), Info 1 (78.55%), and Info 5 (77.57%) show the highest positive shares, and only Info 7 tips slightly negative, at 52.09% (Fig. 5b).

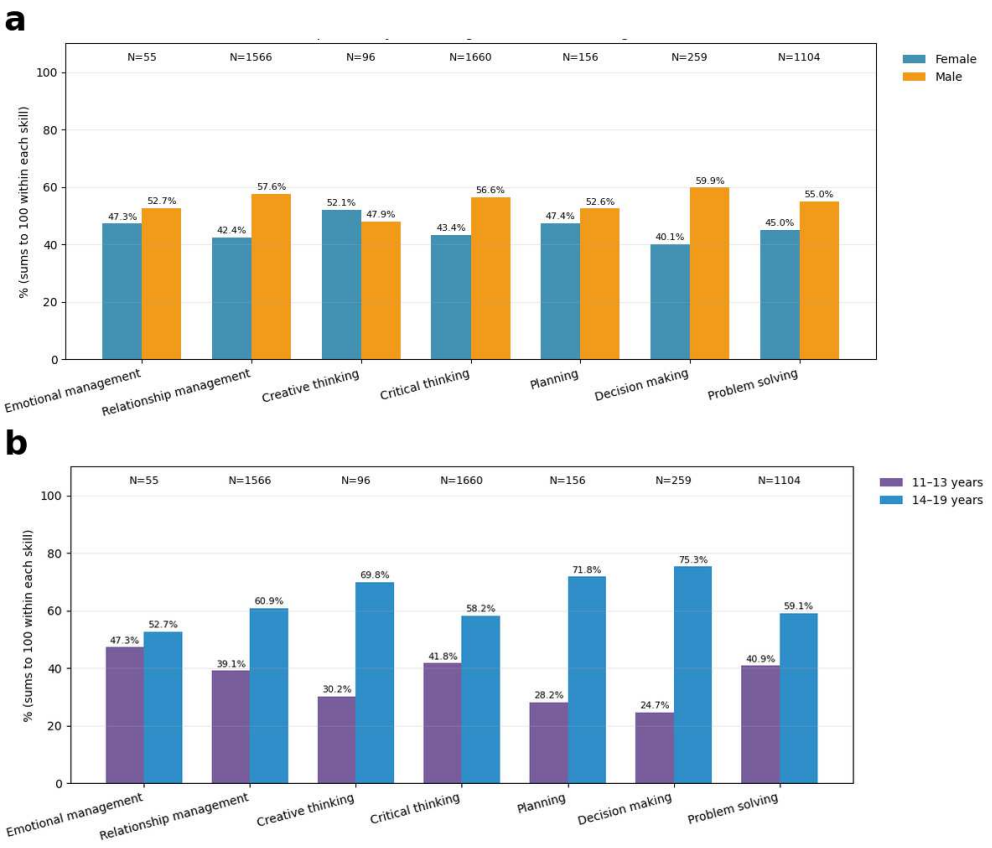

Fig. 2 Skills activated in responses on peers' views (D1). a, Distribution by sex. **b**, Distribution by age group. Percentages sum to 100 within each skill; N above each pair of bars is the number of documents in which the skill was activated.

D3 is more heterogeneous, with three Info units showing negative prevalence, namely Info 7 (71.25%), Info 3 (63.34%), and Info 8 (56.74%). Info 10 (85.44%), Info 5 (77.26%), and Info 1 (64.10%) remain strongly positive (Fig. 5c). This greater heterogeneity fits the agentive nature of D3, which asks respondents to evaluate their own capacity for change rather than describe states or external expectations.

3.2.3 Content categories of the representations

Content category analysis focused on six Info units selected as reference cases on the basis of their content (Info 2, 6, 11, 4, 8, and 9), with category distributions compared by sex and school level.

For D1, Info 2, 6 and 11 each trace back to a single dominant category, with limited sex differences and slightly more pronounced school-level differences. For Info 2, these are agreement with strict regulation, the perceived impossibility of limiting marketing, and indifference. For Info 6, they are agreement with limiting consumption to protect health, unwillingness to do so for taste and convenience, and limited interest in the

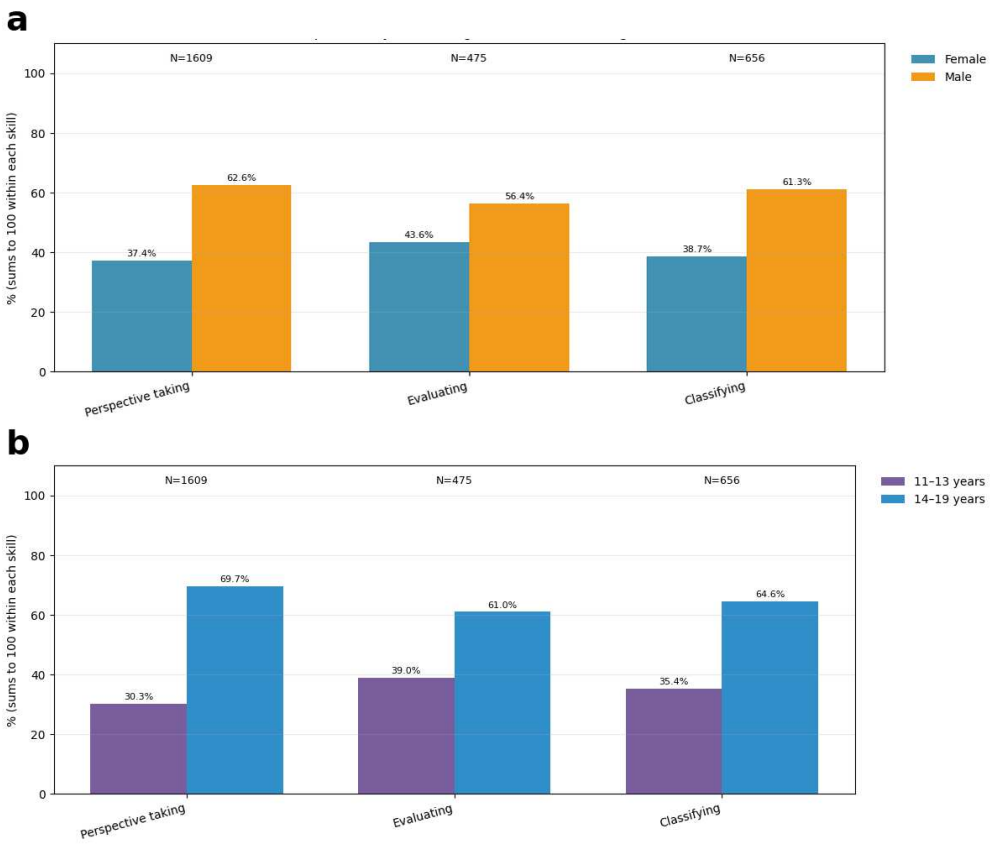

Fig. 3 Skills activated in responses on the role of the environment (D2). a, Distribution by sex. **b**, Distribution by age group. Percentages sum to 100 within each skill; N above each pair of bars is the number of documents in which the skill was activated.

risks. For Info 11, they are the importance of vaccination coupled with a tendency to downplay it, alongside the view that precautions matter more than vaccination itself. Info 4, 8 and 9 show somewhat greater variability and subgroup heterogeneity. Info 8 remains strongly polarised around a single category, the awareness that a behaviour is harmful yet unchanged for social reasons; Info 9 shows a narrower majority and more pronounced gender differentiation, namely agreement on seeking medical advice first, limited awareness of supplement risks, and the belief that supplements improve appearance; and Info 4 follows a similar pattern, with difficulty maintaining a healthy diet, limited awareness of the obesity-cancer link, and a focus on physical appearance.

For D2, a single response category dominates across all six units, more markedly than for D1, with only slight fluctuations by school level and no reversal of the ordering. For Info 2, 6 and 11, these are raising awareness, educating, and informing alongside limiting junk food for Info 2; awareness campaigns and school-family cooperation on healthy snacks and home-cooked food for Info 6; and more information on infection, peer discussion groups, and expanded HPV vaccination campaigns

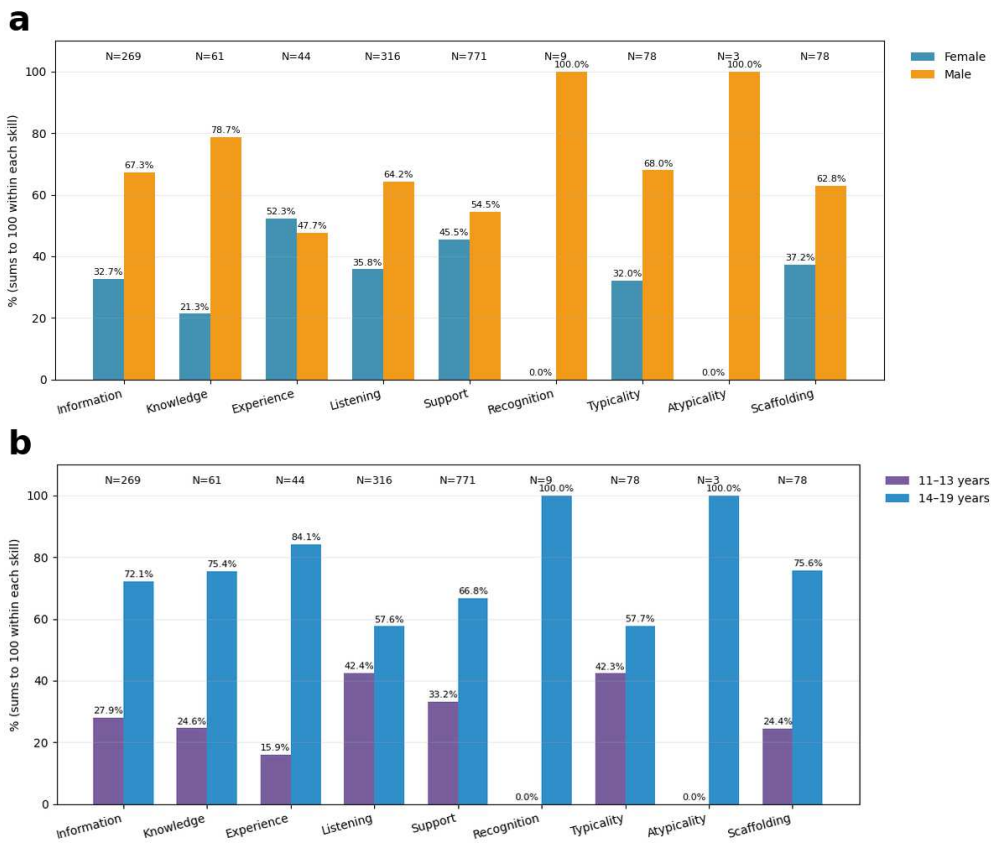

Fig. 4 Needs expressed in responses on personal action (D3). a, Distribution by sex. **b**, Distribution by age group. Percentages sum to 100 within each need; N above each pair of bars is the number of documents in which the need was expressed.

for Info 11. For Info 4, 8 and 9, they are the school promoting good practice, greater dietary information, and family education for Info 4; more information on risks, reduced consumption, and limits on sale for Info 8; and more information on supplement use, monitoring use, and encouraging medically supervised consumption for Info 9. Variability here is attributable mainly to each Info unit's content rather than to sociodemographic factors.

For D3, concentration is intermediate. Info 6 shows a fairly balanced split between conscious dietary choices and limiting processed foods. Info 4, 8 and 9 show sharper concentration, often exceeding 50% for a single category, namely healthy diet and physical activity for Info 4, limiting sugary and alcoholic beverages for Info 8, and consuming only when needed for Info 9. Sex differences remain moderate, while school-level differences, though detectable, do not alter the category ordering.



Fig. 5 Positive and negative sentiment across the 12 Info units. Shares of positive and negative emotion associations (NRC Emotion Lexicon) in the responses to each unit. **a**, Responses on peers' views (D1). **b**, Responses on the role of the environment (D2). **c**, Responses on personal action (D3). Green bars, positive share; red bars, negative share. Units are ordered by decreasing negative share. Info units: 1, healthy lifestyles and cancer prevention; 2, marketing; 3, smoking; 4, diet and body weight; 5, physical activity; 6, processed foods; 7, red meat; 8, sugary and alcoholic beverages; 9, dietary supplements; 10, sun protection; 11, HPV vaccination; 12, online health information.

3.2.4 Pre-post intervention thematic shifts (LDA)

Results of the LDA analysis are presented separately for each of the three questions in Tables 2–4, applying the robustness criterion described above, namely a 95 percent bootstrap confidence interval excluding zero in both the unweighted and length-weighted panels.

D1, peers' views. Of the ten topics, only one shows a robust post-intervention shift by the stated criterion (Table 2). Topic 3, with representative keywords *nulla*,

importa, *fregano*, and *ignorano*, relating to indifference, increased, with an unweighted Δ of $+0.047$ (95% CI $[0.020, 0.075]$) and a length-weighted Δ of $+0.033$ (95% CI $[0.001, 0.067]$). Two further topics reached the criterion in one panel only and are therefore not treated as robust. Topic 8, with keywords *consapevoli*, *rischi*, and *importanza*, decreased in the unweighted panel but not the weighted one, while Topic 7, with keywords *corrette*, *stile*, *indicazioni*, and *veritiere*, increased in the weighted panel only.

D2, role of the environment. Two topics show robust shifts in both panels (Table 3). Topic 4, with keywords *scuola*, *famiglia*, *centri*, and *ambiente*, decreased, with an unweighted Δ of -0.031 (95% CI $[-0.045, -0.017]$) and a weighted Δ of -0.091 (95% CI $[-0.145, -0.033]$), while Topic 7, with keywords *sensibilizzare*, *educare*, and *controllare*, increased, with an unweighted Δ of $+0.039$ (95% CI $[0.017, 0.063]$) and a weighted Δ of $+0.035$ (95% CI $[0.014, 0.059]$). Three further topics reached the criterion in one panel only and are therefore not treated as robust: topics 10 and 8 in the unweighted panel and topic 9 in the length-weighted panel.

D3, personal action. One topic shows a robust shift (Table 4). Topic 5, with keywords *vita*, *stile*, *salute*, *scelte*, and *consapevoli*, decreased, with an unweighted Δ of -0.023 (95% CI $[-0.039, -0.007]$) and a weighted Δ of -0.080 (95% CI $[-0.146, -0.013]$). Topic 3, with keywords *boh*, *seguire*, *consigli*, and *genitori*, shows a borderline increase, with the confidence interval touching zero at the lower bound in both panels, unweighted $[0.000, 0.041]$ and weighted $[0.000, 0.032]$, and should therefore be treated as marginal rather than robust.

3.3 Cognitive style and representations of healthy lifestyles

Data obtained through administration of the P.I.A. instrument at baseline were analysed on two independent subsamples, comprising 534 upper secondary school students and 238 lower secondary school students. Within the upper secondary group, the composition by sex proved markedly unbalanced, with 467 males and 67 females, an element that limits the generalisability of the estimates to the female subgroup. Within the lower secondary sample, the distribution by sex was more balanced, with 141 males and 97 females. In classifying the dominant cognitive profile, participants showing a tie between two or more scales were retained within a residual category, labelled Tied, without any forced assignment. The distribution of cognitive styles measured through the P.I.A. instrument is reported in Table 5, separately for the two groups. Within the upper secondary sample, the distribution proved evenly balanced between the analytical mode, L, at 23.8%, and the concrete mode, U, also at 23.8%, followed by the abstract mode, E, at 18.7%, and the inventive mode, C, at 10.5%. Within the lower secondary sample, a slight prevalence of the analytical mode emerged (28.6%), followed by the concrete mode (25.2%), the abstract mode (17.7%), and the inventive mode (12.6%). Comparison between the two groups reveals a broadly stable hierarchical structure of cognitive styles, with the analytical and concrete modes ranking highest in both cases. The one difference is that the analytical mode shows a slight edge within the lower secondary group, whereas among upper secondary students the analytical and concrete modes prove equivalent. The inventive dimension carries only marginal weight within both groups.

Table 2 Post–pre differences in mean topic proportions for responses on peers' views (D1). $\bar{\theta}_{\text{pre}}$ and $\bar{\theta}_{\text{post}}$, mean topic proportions in the pre- and post-intervention samples; Δ_k, post–pre difference for topic k; 95% CI, percentile bootstrap confidence interval (B = 10,000). Panel A, unweighted means; panel B, means weighted by document length. Keywords are reported in Italian as they appear in the corpus.

Panel A — Unweighted means					
Topic	**Keywords**	$\bar{\theta}_{\text{pre}}$	$\bar{\theta}_{\text{post}}$	Δ_k	**95% CI**
topic 1	mangiare, bevande, fumare, salute	0.254	0.225	−0.029	[−0.060, 0.003]
topic 2	conoscenza, informazioni, peso	0.079	0.078	−0.001	[−0.017, 0.017]
topic 3	nulla, importa, fregano, ignorano	0.122	0.170	+0.047	[+0.020, +0.075]
topic 4	prendono, informati, sottovalutano	0.068	0.056	−0.012	[−0.025, 0.002]
topic 5	indicazioni, vita, seguono, sano	0.078	0.076	−0.002	[−0.019, 0.016]
topic 6	prevenire, consapevoli, malattie	0.076	0.067	−0.009	[−0.025, 0.008]
topic 7	corrette, stile, indicazioni, veritiere	0.070	0.080	+0.011	[−0.009, 0.031]
topic 8	consapevoli, rischi, importanza	0.087	0.061	−0.025	[−0.043, −0.007]
topic 9	accordo, aiutare, meglio, persone	0.108	0.125	+0.017	[−0.006, 0.041]
topic 10	dipende, persona, genitori, consigli	0.060	0.063	+0.003	[−0.013, 0.020]
Panel B — Length-weighted means					
Topic	**Keywords**	$\bar{\theta}_{\text{pre}}$	$\bar{\theta}_{\text{post}}$	Δ_k	**95% CI**
topic 1	mangiare, bevande, fumare, salute	0.309	0.288	−0.021	[−0.071, 0.030]
topic 2	conoscenza, informazioni, peso	0.065	0.076	+0.011	[−0.015, 0.044]
topic 3	nulla, importa, fregano, ignorano	0.115	0.148	+0.033	[+0.001, +0.067]
topic 4	prendono, informati, sottovalutano	0.065	0.062	−0.003	[−0.031, 0.028]
topic 5	indicazioni, vita, seguono, sano	0.054	0.067	+0.013	[−0.011, 0.040]
topic 6	prevenire, consapevoli, malattie	0.056	0.042	−0.015	[−0.032, 0.004]
topic 7	corrette, stile, indicazioni, veritiere	0.036	0.059	+0.024	[+0.001, +0.053]
topic 8	consapevoli, rischi, importanza	0.195	0.124	−0.072	[−0.145, 0.006]
topic 9	accordo, aiutare, meglio, persone	0.077	0.083	+0.007	[−0.016, 0.031]
topic 10	dipende, persona, genitori, consigli	0.029	0.051	+0.022	[−0.004, 0.059]

Table 3 Post–pre differences in mean topic proportions for responses on the role of the environment (D2). $\bar{\theta}_{\text{pre}}$ and $\bar{\theta}_{\text{post}}$, mean topic proportions in the pre- and post-intervention samples; Δ_k, post–pre difference for topic k; 95% CI, percentile bootstrap confidence interval (B = 10,000). Panel A, unweighted means; panel B, means weighted by document length. Keywords are reported in Italian as they appear in the corpus.

Panel A — Unweighted means

Topic	**Keywords**	$\bar{\theta}_{\text{pre}}$	$\bar{\theta}_{\text{post}}$	Δ_k	**95% CI**
topic 10	figli, mangiare, sport, evitare	0.202	0.164	−0.038	[−0.066, −0.010]
topic 4	scuola, famiglia, centri, ambiente	0.082	0.051	−0.031	[−0.045, −0.017]
topic 5	promuovere, servizi, educazione	0.069	0.060	−0.009	[−0.024, 0.007]
topic 1	cibi, mangiare, sano, promuovere	0.139	0.133	−0.006	[−0.030, 0.018]
topic 6	ambiente, esempio, stile, favorire	0.093	0.089	−0.004	[−0.024, 0.017]
topic 2	informazione, attività, istruire	0.070	0.072	+0.002	[−0.017, 0.022]
topic 3	mangiare, sport, fumare, bere	0.110	0.119	+0.009	[−0.015, 0.033]
topic 9	parlarne, aiutare, sport, informazioni	0.083	0.099	+0.016	[−0.007, 0.040]
topic 8	informare, spiegare, scuola, giovani	0.077	0.100	+0.023	[+0.003, +0.044]
topic 7	sensibilizzare, educare, controllare	0.073	0.112	+0.039	[+0.017, +0.063]

Panel B — Length-weighted means

Topic	**Keywords**	$\bar{\theta}_{\text{pre}}$	$\bar{\theta}_{\text{post}}$	Δ_k	**95% CI**
topic 10	figli, mangiare, sport, evitare	0.223	0.229	+0.006	[−0.042, 0.053]
topic 4	scuola, famiglia, centri, ambiente	0.180	0.089	−0.091	[−0.145, −0.033]
topic 5	promuovere, servizi, educazione	0.127	0.083	−0.044	[−0.109, 0.013]
topic 1	cibi, mangiare, sano, promuovere	0.116	0.136	+0.021	[−0.011, 0.054]
topic 6	ambiente, esempio, stile, favorire	0.081	0.086	+0.005	[−0.023, 0.035]
topic 2	informazione, attività, istruire	0.044	0.060	+0.016	[−0.007, 0.043]
topic 3	mangiare, sport, fumare, bere	0.080	0.086	+0.006	[−0.017, 0.030]
topic 9	parlarne, aiutare, sport, informazioni	0.047	0.074	+0.027	[+0.001, +0.060]
topic 8	informare, spiegare, scuola, giovani	0.058	0.078	+0.020	[−0.005, 0.043]
topic 7	sensibilizzare, educare, controllare	0.045	0.080	+0.035	[+0.014, +0.059]

Table 4 Post–pre differences in mean topic proportions for responses on personal action (D3). $\bar{\theta}_{\text{pre}}$ and $\bar{\theta}_{\text{post}}$, mean topic proportions in the pre- and post-intervention samples; Δ_k, post–pre difference for topic k; 95% CI, percentile bootstrap confidence interval (B = 10,000). Panel A, unweighted means; panel B, means weighted by document length. Keywords are reported in Italian as they appear in the corpus.

Panel A — Unweighted means

Topic	Keywords	$\bar{\theta}_{\text{pre}}$	$\bar{\theta}_{\text{post}}$	Δ_k	95% CI
topic 5	vita, stile, salute, scelte, consapevoli	0.085	0.062	−0.023	[−0.039, −0.007]
topic 7	ascoltare, cercare, sport, evitare	0.096	0.082	−0.014	[−0.033, 0.007]
topic 9	salute, capire, suggerimenti, rischi	0.079	0.071	−0.008	[−0.025, 0.009]
topic 4	attenzione, alimentazione, sana, sport	0.095	0.089	−0.006	[−0.027, 0.015]
topic 1	bevande, cibi, consumare, zuccherate	0.062	0.064	+0.002	[−0.014, 0.018]
topic 8	saprei, attenzione, persona, cambiare	0.074	0.078	+0.005	[−0.013, 0.024]
topic 2	mangiare, fumare, bere, sano, sport	0.273	0.281	+0.008	[−0.028, 0.046]
topic 10	cercare, fumare, smettere, cibo	0.079	0.087	+0.008	[−0.012, 0.030]
topic 6	evitare, dieta, attività, fumare	0.096	0.105	+0.009	[−0.014, 0.032]
topic 3	boh, seguire, consigli, genitori	0.062	0.081	+0.019	[0.000, 0.041]

Panel B — Length-weighted means

Topic	Keywords	$\bar{\theta}_{\text{pre}}$	$\bar{\theta}_{\text{post}}$	Δ_k	95% CI
topic 5	vita, stile, salute, scelte, consapevoli	0.171	0.090	−0.080	[−0.146, −0.013]
topic 7	ascoltare, cercare, sport, evitare	0.094	0.092	−0.003	[−0.035, 0.034]
topic 9	salute, capire, suggerimenti, rischi	0.105	0.080	−0.025	[−0.063, 0.016]
topic 4	attenzione, alimentazione, sana, sport	0.079	0.068	−0.012	[−0.036, 0.015]
topic 1	bevande, cibi, consumare, zuccherate	0.067	0.087	+0.020	[−0.019, 0.067]
topic 8	saprei, attenzione, persona, cambiare	0.042	0.042	+0.000	[−0.015, 0.016]
topic 2	mangiare, fumare, bere, sano, sport	0.271	0.320	+0.049	[−0.007, 0.106]
topic 10	cercare, fumare, smettere, cibo	0.057	0.073	+0.016	[−0.010, 0.045]
topic 6	evitare, dieta, attività, fumare	0.084	0.104	+0.020	[−0.011, 0.052]
topic 3	boh, seguire, consigli, genitori	0.029	0.044	+0.015	[0.000, 0.032]

Table 5 Predominant cognitive style (P.I.A.) by school level. Upper secondary, N = 534; lower secondary, N = 238. Tied, participants with equal scores on two or more styles.

Profile	Upper sec. freq.	Upper sec. %	Lower sec. freq.	Lower sec. %
L (Analytical)	127	23.8	68	28.6
U (Concrete)	127	23.8	60	25.2
C (Inventive)	56	10.5	30	12.6
E (Abstract)	100	18.7	42	17.7
TIED	124	23.2	38	16.0
Total	**534**	**100**	**238**	**100**

To explore whether cognitive style profiles were associated with response patterns across the INFO questions, an additional correlational analysis was conducted at two levels, namely thematic structure, based on per document LDA topic proportions, and emotional content, based on emotion vectors derived from the Italian NRC EmoLex lexicon according to Plutchik's model [29]. In the absence of individual identifiers allowing direct linkage between the P.I.A. file and the INFO files, the connection between the two datasets was constructed through random sampling with replacement, on the assumption that the two samples were exchangeable, as both were drawn from the same school population. For each bootstrap iteration, with B set to 100, N subjects were sampled with replacement from the P.I.A. file and N documents from the INFO files, and the Spearman correlation was computed between the continuous scores on each psychological scale and the corresponding textual scores. The structure of this analysis is summarised in Table 6.

Table 6 Structure of the exploratory bootstrap correlation analysis between P.I.A. scores and INFO textual indicators. L, analytical; U, concrete; C, inventive; E, abstract style. LDA, Latent Dirichlet Allocation topic proportions; NRC, emotion scores from the NRC Emotion Lexicon.

School level	Tool	Scale	Textual target	N correlations
Upper secondary	PIA	L, U, C, E	10 topics LDA × 3 questions	120
Lower secondary	PIA	L, U, C, E	10 topics LDA × 3 questions	120
			Total LDA	**240**
Upper secondary	PIA	L, U, C, E	8 emotions NRC × 3 questions	96
Lower secondary	PIA	L, U, C, E	8 emotions NRC × 3 questions	96
			Total NRC	**192**

Of a total of 240 correlations involving the LDA topics and 192 involving the NRC emotions, none proved robust according to the adopted criterion, namely a bootstrap 95% confidence interval excluding zero. It should be noted that, given the exploratory nature of this analysis, these results were not corrected for multiple comparisons. It should also be emphasised that, in the absence of genuine individual-level linkage

between the P.I.A. file and the INFO files, this procedure randomly pairs psychometric scores and textual documents belonging to different individuals. A correlation constructed in this way tends, by design, toward the null value, regardless of whether a genuine association exists at the individual level. The result obtained cannot therefore be interpreted as evidence supporting the rejection of H2, but should instead be read as an indication of the absence of a detectable association, at the aggregate level, between the distributions of the two samples. A more rigorous test of H2 would require the collection of data including individual identifiers that would allow direct linkage between psychometric measures and the textual responses of the same student.

3.4 Students' Human Satisfaction

The final results reported here come from the "Human Satisfaction" questionnaire, administered at the end of the project and completed by 398 students, 89.7% of whom were Italian. Because this brief three-item questionnaire was administered independently of the INFO post-test and completion was not conditional on completing the INFO instrument, its respondents only partially overlap with the post-intervention INFO sample (N = 384), and the two samples differ in composition. Of the 398 respondents, 217 were aged 11 to 13 and 181 were aged 14 to 19, while 247 were male and 151 were female. As noted, the questionnaire assessed students' satisfaction with their participation in the project and their intention, following the intervention, to change a personal behaviour that the project had helped them recognise as unhealthy. Overall, 82% of respondents rated the project as useful and 63% expressed the intention to change their own behaviour in a healthier direction after taking part in the project.

4 Discussion

The present study analyzed changes associated with a school-based intervention program for cancer prevention aimed at moving beyond the mere transmission of information, in order to transform the social and individual representations of adolescents regarding healthy lifestyles [15, 33, 34]. The data are consistent with previous studies showing that, even when adolescents possess good theoretical knowledge of cancer risk factors, this knowledge does not automatically translate into healthy behaviors [12, 13, 21, 34, 35]. In the present sample, this gap is reflected in the weak association between knowledge and adherence to healthy lifestyles observed in the present sample (Section 3.1). Health information therefore appears necessary but not sufficient, particularly among males and older adolescents, in whom the link between knowing and acting appears weaker.

The pre–post comparisons of the INFO narratives are partly consistent with H1. The clearest pattern concerns the role attributed to the environment: in the post-intervention sample, the topic attributing responsibility to school, family and other institutions was less frequent, whereas the topic centred on raising awareness, educating and monitoring was more frequent. This shift from a frame of "institutional delegation" to one of "direct educational engagement" is compatible with adolescents coming to view prevention as an active process in which they themselves can act as promoters [21, 35, 36]. In responses on personal action, the topic expressing generic

references to lifestyle, health and conscious choices became less frequent; this may indicate a move away from abstract notions of a healthy lifestyle, a dimension of agency that is relevant to adolescent health promotion [34, 37], but no topic referring to specific actions increased robustly, and the marginal increase of a topic expressing uncertainty and reliance on parents' advice calls for caution. Finally, the increase in references to peers' indifference may reflect a more realistic perception of the resistance present within adolescents' own social group, although it may also reflect genuine disengagement. Because only four of 30 topic-level comparisons met the robustness criterion and the two samples could not be linked, these shifts should be regarded as exploratory.

The Human Satisfaction data are consistent with these findings: 82% of adolescents considered the project useful and, more importantly from a public health standpoint, 63% expressed an intention to change unhealthy behaviors, although these data are self-reported and not compared with a control group. Moreover, the Service Learning approach [26, 27], implemented through the design of community campaigns, is a plausible mechanism for engaging students as health promoters rather than passive recipients, although this mechanism was not tested directly.

The sentiment analysis revealed further noteworthy aspects concerning how adolescents experienced engagement with sensitive cancer-prevention issues through health-related choices. The data showed that not all prevention themes are received in the same way. While messages on sun protection, physical activity and healthy lifestyles in general elicit a predominantly positive sentiment, topics such as red meat consumption, smoking and sugary or alcoholic beverages elicit more critical or negative emotional responses. These themes concern behaviors embedded in daily eating habits and adolescent sociality, which may be more resistant to change and may require interventions that take into account emotional barriers and peer pressure.

The value of the INFO narrative tool should also be emphasized, as it enabled the detection of both representations and competencies (life and soft skills) activated during reflection on health and cancer-prevention themes. When reflecting on peer attitudes toward health-related requests and guidance, the competencies that emerged most prominently were critical thinking, relationship management and problem solving, suggesting that the tool can activate complex mental processes and active evaluation. With regard to the role of the environment (family, school, healthcare system), the soft skills most frequently elicited were perspective taking, classification, and evaluation. This suggests that adolescents recognize the complexity of the social system to which they belong and are able to analyze the responsibilities of its different actors.

Concerning the emerging needs of adolescents with respect to lifestyle change, elicited by the third question in the INFO instrument, support was the most frequently expressed need. This suggests that, despite adolescents possessing high levels of theoretical knowledge (84.9% starting from a high baseline level), they feel the need to be accompanied and supported in the critical transition from "knowing" to "doing." Documents expressing needs for information and knowledge were predominantly written by males, whereas experience was the only need with a slight female majority. Finally, the emergence of needs related to being listened to and to scaffolding suggests that

adolescents perceive health as a path requiring guidance, orientation and support that is relational as well as technical in nature.

Although the study hypothesized a correlation between cognitive styles and representations, the aggregate analyses did not reveal robust associations; as noted, however, this analysis could not test H2 at the individual level. The prevalence of analytical and concrete styles within the sample, which remained substantially stable across age groups, nonetheless suggests that adolescents tend to approach problems either through the logical decomposition of data or through practical, immediate experience. Nevertheless, cognitive profiling provides important indications: the fact that approximately half of the students showed a predominantly analytical or concrete style suggests that prevention messages should be dual, combining rigorous scientific data with practical experiences and immediate benefits. Furthermore, strengthening the analytical style through critical thinking and planning may support the executive functions associated with the prevention of chronic disease. In summary, the combination of logical reasoning and practical anchoring may represent a useful foundation for schools seeking to transform knowledge into agency.

"I Care for Me for You" thus represents a promising model of school-based cancer prevention intervention in the Italian context; a model based on an approach that does not merely transmit information to promote knowledge, but aims to transform the representations held by adolescents regarding health and how to manage it. Information alone does not appear sufficient; although adolescents frequently "know" what is harmful, their needs for support and for being listened to suggest that change may depend on emotional and relational variables, as well as on the capacity of the environment to activate individual agency.

Among the principal strengths of the study, two may be highlighted: on the one hand, the innovative methodological approach, and on the other, the nature of the INFO instrument used to assess changes in representations. With regard to the innovative methodological approach, the use of Service Learning and Life Skills allowed students to move beyond the role of passive recipients, becoming active "health promoters" engaged in developing campaigns for their own communities; this also defines an important community-level impact, extending the reach of the project beyond school walls and raising awareness among families and local governance bodies. The project suggests that fostering autonomy and responsibility among young people may help reduce the gap between theoretical knowledge and practical behavior, a gap often reinforced by peer pressure and misleading marketing.

With respect to the nature of the narrative instrument "INFO," it should be noted that this tool served a dual function: not only as an assessment of representations, but also as an intervention in itself, stimulating critical thinking and perspective taking during completion. A further strength lies in the integration between primary and secondary prevention.

This study has several limitations. The design had no control group, so the observed differences cannot be attributed causally to the intervention. Because data were collected anonymously, pre- and post-intervention responses could not be linked, and comparisons are between independent samples. The post-intervention sample was much smaller than the pre-intervention one and included a higher proportion of males,

so differences in composition may contribute to the observed shifts. Students were nested within classes and schools, a structure that was not modelled, and confidence intervals may therefore underestimate uncertainty. The narrative instrument and the knowledge, adherence and satisfaction questionnaires have not been formally validated, the keyword-based coding was not assessed for inter-rater agreement, and only four of 30 topic-level comparisons met the robustness criterion without correction for multiple testing. The distributions of skills and needs by sex and school level describe the composition of the documents rather than rates per student, and the bootstrap pairing of P.I.A. and INFO data cannot test associations at the individual level.

The study suggests the importance of incorporating the project into the Extended Educational Offer Plan (POF) of schools, thereby enabling cancer prevention to become a stable curricular practice rather than an isolated event.

It further highlights the need for a life-course approach to prevention: intervening during adolescence, identified as a "window of susceptibility", as well as during other sensitive developmental stages, in order to consolidate protective habits capable of persisting across the entire life course.

Acknowledgements. We thank the students, teachers and administrators of the participating schools and the professionals of the Oncology Network of Piedmont and Aosta Valley. We also thank the Italian Society of Pediatric Psychology (SIPPed) for its support in defining the monitoring and evaluation plan.

Declarations

Funding. This work was supported by the CRT Foundation (Fondazione Cassa di Risparmio di Torino).

Competing interests. The authors declare no competing interests.

Data availability. Anonymised data are available from the corresponding author upon reasonable request.

Materials availability. The questionnaires and the INFO narrative instrument are available from the corresponding author upon reasonable request.

Author contributions. F.C., L.M., A.C. and C.Pi. implemented the intervention and carried out data collection and data management. M.Sp. and M.Sc. designed the statistical and methodological approach and performed the data analysis. F.C., C.Po., G.P. and F.A. defined the study objectives, interpreted the results and wrote the first draft of the manuscript. All authors reviewed and approved the final manuscript.

References

[1] World Health Organization. Global cancer burden growing, amidst mounting need for services. World Health Organization news release (2022).

[2] Associazione Italiana di Oncologia Medica. *I numeri del cancro in Italia 2024 [Cancer figures in Italy 2024]* (Intermedia Editore, 2024).

[3] American Cancer Society. *Cancer Facts & Figures 2024* (American Cancer Society, 2024).

[4] World Cancer Research Fund. Diet, nutrition, physical activity and cancer: A global perspective. Third Expert Report, World Cancer Research Fund International (2021).

[5] Kohler, L. N. *et al.* Adherence to the cancer prevention guidelines and cancer incidence, mortality, and all-cause mortality: A systematic review and meta-analysis. *Cancer Epidemiology, Biomarkers & Prevention* **25**, 1018–1028 (2016).

[6] European Commission. Europe's Beating Cancer Plan (2021). URL https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021DC0044. COM(2021) 44 final.

[7] Santelli, J. S., Smith Rogers, A. & Rosenfeld, W. Adolescent health and disease prevention: A lifelong impact. *Journal of Adolescent Health* **52**, S1–S7 (2013).

[8] Parajuli, J. & Prangthip, P. Adolescent nutrition and health: a critical period for nutritional intervention to prevent long term health consequences. *Current Nutrition Reports* **14**, 116 (2025).

[9] Fuemmeler, B. F., Pendzich, M. K. & Tercyak, K. P. Weight, dietary behavior, and physical activity in childhood and adolescence: implications for adult cancer risk. *Obesity Facts* **2**, 179–186 (2009).

[10] White, M. C. *et al.* Age and cancer risk: a potentially modifiable relationship. *American Journal of Preventive Medicine* **46**, S7–S15 (2014).

[11] Guthold, R., Stevens, G. A., Riley, L. M. & Bull, F. C. Global trends in insufficient physical activity among adolescents: A pooled analysis of 298 population-based surveys with 1.6 million participants. *The Lancet Child & Adolescent Health* **4**, 23–35 (2020).

[12] Di Giuseppe, G., Pelullo, C. P., Lanzano, R., Napolitano, F. & Angelillo, I. F. Knowledge, attitudes, and behaviors towards cancer prevention among adolescents in Italy. *BMC Public Health* **20**, 1592 (2020).

[13] Silva, M. & Ferreira, C. Aesthetic ideals versus health choice in adolescents: A qualitative study. *Appetite* **138**, 201–209 (2019).

[14] Valanju, R. *et al.* Youth perspective on chronic disease prevention. *The Lancet Child & Adolescent Health* **6**, e13–e14 (2022).

[15] Pugh, G., Hough, R. E., Gravestock, H. L., Jackson, S. E. & Fisher, A. The health behavior information needs and preferences of teenage and young adult cancer survivors. *Journal of Adolescent and Young Adult Oncology* **6**, 318–326 (2017).

[16] Al-Azri, M., Al-Rubaie, K., Al-Nabhani, M., Al-Busaidi, M. & Davidson, R. Knowledge, attitudes, and barriers towards cancer prevention and early detection among adolescents: A cross-sectional study. *BMC Public Health* **19**, 1321 (2019).

[17] Kyle, R. G., Forbat, L. & Hubbard, G. Cancer awareness among adolescents in Britain: a cross-sectional study. *BMC Public Health* **12**, 580 (2012).

[18] Hall, P. A., Fong, G. T. & Epp, L. J. Cognitive and personality factors in the prediction of health behaviors: an examination of total, direct and indirect effects. *Journal of Behavioral Medicine* **37**, 1057–1068 (2014).

[19] Del Missier, F., Mäntylä, T. & de Bruin, W. B. Decision-making competence, executive functioning, and intelligence in young and older adults. *Journal of Behavioral Decision Making* **25**, 331–351 (2012).

[20] Sharma, L., Markon, K. E. & Clark, L. A. Toward a theory of distinct types of "impulsive" behaviors: A meta-analysis of self-report and behavioral measures. *Psychological Bulletin* **140**, 374–408 (2014).

[21] McGovern, C. M., Militello, L. K., Arcoleo, K. J. & Melnyk, B. M. Factors associated with healthy lifestyle behaviors among adolescents. *Journal of Pediatric Health Care* **32**, 473–480 (2018).

[22] Perricone, G. & Polizzi, C. *Valutare gli stili cognitivi nel problem solving: Prove di Intelligenze Applicate [Assessing cognitive styles in problem solving: Applied Intelligence Tests]* (Erickson, Trento, 2010).

[23] Cecchini, M. *Processi cognitivi e sistemi di sicurezza nello sviluppo infantile [Cognitive processes and safety systems in child development]* (La Nuova Italia Scientifica, 1987).

[24] World Health Organization. Life skills education for children and adolescents in schools. Tech. Rep. WHO/MNH/PSF/93.7A.Rev.1, World Health Organization, Geneva (1994).

[25] Kechagias, K. (ed.) *Teaching and Assessing Soft Skills* (1st Second Chance School of Thessaloniki (Neapolis), MASS Project, Thessaloniki, 2011).

[26] Tapia, M. N. *Il Service-Learning nella scuola: Teorie e pratiche di solidarietà e apprendimento [Service-Learning in schools: Theories and practices of solidarity and learning]* (Città Nuova, 2026).

[27] Fiorin, I. (ed.) *Oltre l'aula: La proposta pedagogica del Service Learning [Beyond the classroom: The pedagogical proposal of Service Learning]* (Mondadori Università, 2016).

[28] World Health Organization. *WHO Report on Cancer: Setting Priorities, Investing Wisely and Providing Care for All* (World Health Organization, Geneva, 2021).

[29] Plutchik, R. in *A general psychoevolutionary theory of emotion* (eds Plutchik, R. & Kellerman, H.) *Emotion: Theory, Research, and Experience. Vol. 1: Theories of Emotion* 3–33 (Academic Press, New York, 1980).

[30] Plutchik, R. The nature of emotions. *American Scientist* **89**, 344–350 (2001).

[31] Mohammad, S. M. & Turney, P. D. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence* **29**, 436–465 (2013).

[32] Blei, D. M., Ng, A. Y. & Jordan, M. I. Latent Dirichlet allocation. *Journal of Machine Learning Research* **3**, 993–1022 (2003).

[33] Kyle, R. G. *et al.* Scottish adolescents' sun-related behaviours, tanning attitudes and associations with skin cancer awareness: A cross-sectional study. *BMJ Open* **4**, e005137 (2014).

[34] de Sousa, Z. A. A., da Silva, J. G. & de Assunção Ferreira, M. Saberes e práticas de adolescentes sobre saúde: Implicações para o estilo de vida e cuidado de si. *Escola Anna Nery Revista de Enfermagem* **18**, 400–406 (2014).

[35] Moreno, L. A. *et al.* Assessing, understanding and modifying nutritional status, eating habits and physical activity in European adolescents: the HELENA (Healthy Lifestyle in Europe by Nutrition in Adolescence) study. *Public Health Nutrition* **11**, 288–299 (2008).

[36] Abraham, O., Rosenberger, C. A., LeMay, S. M. & Bittner, S. J. Adolescents' perceptions about cancer and preferences for cancer education. *Cancer Control* **28**, 10732748211036057 (2021).

[37] Boonekamp, G. M. M., Dierx, J. A. J. & Jansen, E. Shaping physical activity through facilitating student agency in secondary schools in the Netherlands. *International Journal of Environmental Research and Public Health* **19**, 9028 (2022).