

Nature of heat inventory: Assessing students' understanding of thermal processes in kinetic molecular terms

Alexander Volfson,¹ Orly Lankry,² and Yuval Ben Abu^{3,4,5}

¹*Sami Shamoon College of Engineering, Beer Sheva, Israel*

²*Talpiot Academic College for Teachers, Holon, Israel*

³*Physics Department, Tel Aviv University, Israel*

⁴*Sapir Academic College, Sderot, Israel*

⁵*Orot Israel College, Rehovot, 7644709, Israel*



(Received 28 September 2025; revised 5 June 2026; accepted 16 April 2026; published 5 August 2026)

The present study describes the development and field testing of the *nature of heat inventory* (NOHI) that examines high school and undergraduate students' understanding of heat transfer processes in the terms of kinetic molecular theory. NOHI comprises 10 items. Based on preliminary study and the history of science, the 30 distractors of NOHI were designed to reflect the following three categories of possible misconceptions regarding heat: (a) *materialistic views*, (b) *hot atoms*, and (c) *expanding atoms*. Field testing, carried out on 81 students for science teaching, showed the materialistic misconceptions of heat still to be the most popular. NOHI has been proven to be a valid and reliable tool that could be used to determine students' thinking regarding classical heat processes.

DOI: [10.1103/q4n2-d2w3](https://doi.org/10.1103/q4n2-d2w3)

I. INTRODUCTION

If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generations of creatures, what statement would contain the most information in the fewest words? I believe it is the *atomic hypothesis* ... that *all things are made of atoms—little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another*. In that one sentence, you will see, there is an *enormous* amount of information about the world, if just a little imagination and thinking are applied.

This is how Richard Feynman started his seminal course of lectures [1]. This is why the kinetic molecular theory of matter is taught in the very beginning of almost every middle-school science course (physics, chemistry, biology). An important part of this theory is the explanation of basic heat processes: heat transfer, convection, melting,

solidification, thermal expansion, etc. Thus, students finishing middle and furthermore, high school should understand most of these issues and be able to explain simple everyday heat phenomena at least qualitatively. This, however, is not the situation; not only among middle-school students but also among physics and engineering graduates, as our field experience shows. Thus, our experience meets the findings of professional publications of the last decades [2–5].

Two courses of action are needed to improve the situation. The first is diagnostic in character, while the second is practical or curricular in character [6]. Regarding the diagnostic one, the general challenge involves identifying students' conceptual understanding, which will enable teachers to further design more efficient learning environments addressing students' misconceptions on the subject learned and help them reshape their initial knowledge into a more scientifically accepted understanding [7]. In addition, keeping in mind that assessment drives teaching [8], or, in other words, “what gets assessed is what gets taught” [9], the importance of such instruments is obvious. One way to identify students' knowledge about scientific concepts is the use of diagnostic tools—questionnaires, tests, inventories, etc. [6,10–12]. These authors suggest that more tools need to be developed to allow instructors to better evaluate their students' understanding of physics concepts.

Herein, we develop the NOHI measuring students' understanding of basic heat phenomena in qualitative terms of classical kinetic molecular theory as shown in

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

the Appendix. It is applicable in high schools as well as in colleges and universities. NOHI contains 10 items and is designed to be completed within 10–20 min, thus ensuring a high probability of cooperation by the examinee [13]. Each item relates to the same four mental models of heat as we further explain, which should provide some reliable picture of students' thinking.

II. WHY DEVELOP NOHI?

Kinetic molecular theory understanding is crucial for further understanding of physics, chemistry, and other nature sciences. Nevertheless, kinetic molecular theory is a part of almost every physics curriculum; undergraduate students as well as graduate engineers still experience difficulties applying it to explain simple heat processes [14,15] (see also Fig. 1).

This gloomy assessment is not intended as a wholesale indictment of the many dedicated and competent physics teachers. It does tell us, though, that effective instruction requires more than dedication and subject knowledge. It requires technical knowledge about how students think and learn [12].

In other words, more diagnostic instruments are needed to facilitate teachers' probing and assessing their students' concepts and misconceptions [6,10,11]. Several diagnostic instruments on the topic of heat exist in the literature. However, most of them do not probe the basic qualitative understanding of kinetic molecular theory applied in analyzing simple heat phenomena. Indeed, the *heat and temperature conceptual evaluation* developed by Thorton and Sokoloff [16] probes students' understanding in relatively advanced topics such as temperature, specific

heat capacity, change of phase, etc. [4]. Ding *et al.* [17]—energy assessment tool—probe students' concepts about heat in abstract terms of energy. Thermal concept evaluation questionnaire developed by Yeo and Zadnik [18] focuses on thermal processes from the viewpoint of temperature, that is, also energy in other terms. No inventory deals explicitly with the basic inner mechanism of particles' (molecules, atoms, electrons) motion and collisions. The present study aims at filling this gap. NOHI probes examinee's understanding of heat propagation, thermal expansion and compression, and melting and freezing in terms of kinetic molecular theory (see Table I below). It does not use abstract (energy, temperature, and entropy) or advanced (quantum mechanics) terms.

III. STUDENTS' CONCEPTS AND MISCONCEPTIONS OF HEAT

The development of NOHI is based on a preliminary qualitative study including (a) literature review, (b) history of science (HOS) review, (c) clinical, and (d) focus-group interviews [6,10,19]. This preliminary research focused on the conceptual understanding of basic heat processes as listed in Table I. There are two main approaches to students' misconceptions. The shift category approach [2,20] and the framework theory approach (Vosniadou and Brewer [7,21]) both necessitate that there is some degree of coherency in students' views. Indeed, according to Chi, students' knowledge is organized in different ontological categories. Once categorized, a concept can “inherit” features and attributes from its category membership that enable a learner to make many inferences and attributions about a novel concept and/or phenomenon [2,22]. To be more precise, Chi considers *two* categories with which scientific phenomena may be associated:

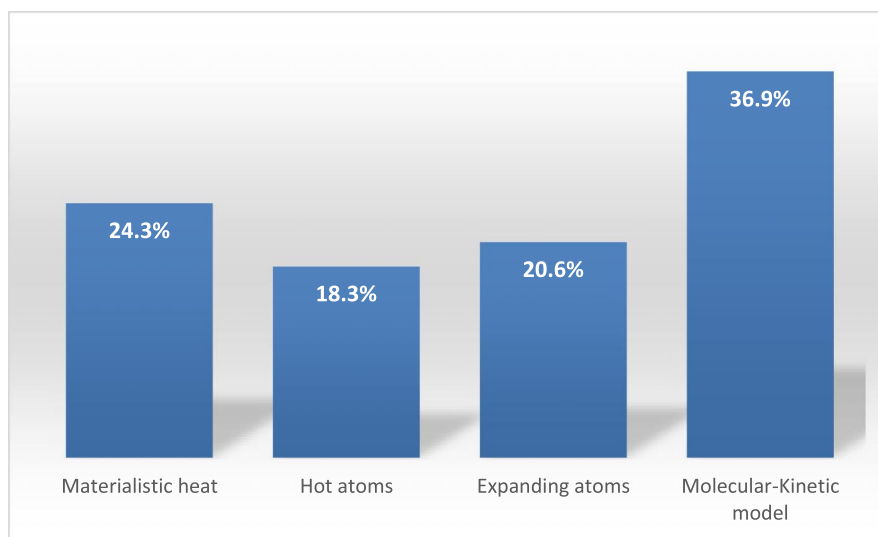


FIG. 1. Distribution of answers among the categories.

TABLE I. Seven basic heat phenomena measured by NOHI.

Phenomenon		Items on NOHI
Heat propagation	<i>Conduction</i>	1, 5, 6
	<i>Convection</i>	2
	<i>Radiation</i>	1
Thermal expansion of liquids and metals		4, 8
Thermal compression of polymers (rubber)		9
Melting		3
Freezing		10

entities and *processes*. Processes can be further distinguished as *sequential* or *emergent*. In sequential processes, each of the system's elements plays a distinct role: the behavior of elements is predefined and predictable; there is an order in the interactions between the elements. In emergent processes, the elements do not act in a predetermined way; there is no order of interactions between elements, and a random number of interactions occur simultaneously. Incorrectly conceiving an emergent process as if it were an entity (matter) or a direct process, according to Chi, is one of the sources for students' robust scientific misconceptions regarding process phenomena such as heat [3]. In this way, a miscategorized concept will often act as a relatively consistent incorrect science theory (alternative theory and mistaken mental model). On the other hand, DiSessa's "pieces in knowledge" theory views naïve knowledge as being fragmented and loosely coherent [23]. In our case, it is expected that answers provided by participants having backgrounds in the physics of heat will have a more degree of coherency, and therefore, we considered the theoretical framework of Chi and Vosniadou to be more appropriate for the present research. Even more so, our field experience and the preliminary study results are both consistent with this approach.

Heat processes, according to Chi [2,3], are classified as emergent processes. In such processes, the elements act randomly (random motion of atoms or molecules in gas, conduction electrons in metals), there is no order of interactions between elements (all the particles' collisions are random), and a random number of interactions (collisions) occur simultaneously. Thus, the resulting global pattern in an emergent process, such as heat conduction, is actually a statistical result of the random interactions between the elements, so ordered-like (or disordered) patterns can emerge from a mess. Two main alternative models for heat transfer were outlined by Chi: materialistic view of heat and considering heat as a sequential (ordered) process. Our preliminary study results reinforce these findings of Chi and bring a new, absent in the professional literature. Actually, we identified five main categories of mental models [2,7,24] regarding heat in the preliminary study:

1. **Heat particles**—naïve students tend to explain phenomena related to heat in terms of matter and the characteristics they attribute to matter. For instance, students mistakenly view heat as being a kind of matter that propagates from one place to another and can be contained and accumulated in a closed space [25]. Such materialistic views of heat were also found among undergraduate students [15].
2. **Heat is an invisible liquid (caloric)**—heat is transferred from hotter to colder bodies by a flow of a special heat liquid called caloric. Caloric can also be contained and accumulated in bodies. This model was popular in the science of XVII—XVIII centuries [26].
3. **Hot atoms or molecules**—when students learn about heat in the terms of air molecules' motion, some of them tend to perceive heat transfer as a direct process of matter propagation according to which "hot molecules" travel from one location to another [20]. This is obviously a hybrid model in terms of Vosniadou (2008) connecting the materialistic view of heat and the idea that heat propagates in the air in terms of the air molecules' motion.
4. **Expanding atoms or molecules**—when we heat a substance (metal rod, water, air in a balloon), it expands. This is due to the increase in volume of every atom and molecule of the substance. Atoms or molecules of a cold substance, on the contrary, usually shrink. For instance, hot air molecules are bloated, therefore, less dense and thus flow up (convection); when we heat an edge of a metal rod, the atoms expand and hit each other along the rod, thus transferring heat. To the best of our knowledge, this way of thinking is absent in the professional literature. However, this idea was more than once repeated in the interviews and further gained certain popularity in the first and second versions of NOHI.
5. **The Kinetic molecular model**—scientifically correct explanations.

Further, while running the first version of NOHI, we saw that participants almost did not choose the *caloric* model, and thus, we merged distractors of type 1 and 2 into one category named **Materialistic heat**. That is, heat is a special matter (particles or liquid) that is transitional—flow from point "A to point B"; containable—can be contained and accumulated and; stable—do not spontaneously appear or disappear [25,27].

IV. DEVELOPING THE INSTRUMENT

A. Methodology used for developing NOHI

We based the development of NOHI mainly on a summary of the methodology presented by Adams and Weiman [13]. We also followed methods used in previous

similar studies developing diagnostic instruments [6,10,17,28]. Adams and Weiman [13] summarized the following four phases of the development process of diagnostic instruments:

- (i) Delineation of the purpose of the test and the scope of the construct or the extent of the domain to be measured;
- (ii) Development and evaluation of the test specifications: item format, time restrictions, characteristics of the population, and test procedures;
- (iii) Development, field testing, evaluation, and selection of the items and scoring guides and procedures;
- (iv) Assembly and evaluation of the test for operational use.

Next, we will discuss every phase in detail.

Phase 1: Delineation of the purpose of the questionnaire and the extent of the domain to be measured

NOHI aims at measuring students' concepts and misconceptions regarding the seven basic heat phenomena in terms of kinetic molecular theory as represented in Table I below. These issues were approved as being sufficient by 12 expert: a physics professor, two Ph. D. in physics, two Ph.D.s in science and technology teaching and three high-school physics teachers, and four M.Sc. physicists.

Phase 2: Development and evaluation of the questionnaire specifications

This phase includes characteristics of the target population, item format decision, time restrictions, and questionnaire procedures (Adams and Weiman [13]).

Characteristics of the target population

The proposed instrument is intended to probe ideas mainly among high school and undergraduate physics, chemistry, and engineering students.

Item format: Student centered

The literature distinguishes between two types of assessment instruments: (a) student-centered, which include in-depth interviews and open-ended questions, and (b) the convergent-type, ones consisting mainly of multiple-choice tasks. Convergent-type instruments, in contrast to student-centered ones, are more objective and demand limited human and time resources [29]. To make the instrument valid, i.e., probe what students really think and not what the authors think that students think, we have to formulate the distractors based on known students' misconceptions [13]. To do so, we grounded the distractors in (a) misconceptions known from professional literature; (b) HOS [10]; (c) preliminary qualitative study in which five are clinical [6,27] and (d) five focus-group interviews [19,30] were held. All the interviews were conducted as semistructured and based on the same ten questions as NOHI (see the Appendix). The focus group interviews were conducted in popular science community events [19] and included about 10 to 50 participants each. These interviews encompassed a multiaged and multieducated

(high school to graduate engineers and physicists) population.

In addition, in order to enable students to articulate whatever ideas they deem representative of their viewpoints on the target issue in an open-ended format, the following choice was added to each of the questionnaire's items: "None of the above choices fits my basic viewpoint. My basic viewpoint is _____ (please explain your viewpoint in the space provided below)" [6,10,31].

Time restrictions

According to Adams and Wieman [13], 30 min is the most acceptable upper limit for a concept questionnaire. An average student could complete NOHI within 10–20 min, as it comes out from the field testing.

Phases 3 and 4: Development, field testing, evaluation, and selection of the items; assembly and evaluation of the questionnaire for operational use

These phases include the methods and data sources for distractors' development, field-testing results, their evaluation, and further fine-tuning of the instrument for operational use [13]. Generally, NOHI went through two field-testing iterations. The first iteration results enabled us to evaluate the reliability and validity of the items, as well as learn about their quality and effectiveness. These helped us fine-tune NOHI, after which we completed the second reiteration that was needed to revise, revalidate, and retune the instrument [6]. We present the results of the second iteration below.

Distractors

To design a student-centered instrument, we used the four categories of misconceptions identified in the preliminary study. In what follows, we illustrate each category by presenting some representative examples of distractors.

1. **Materialistic heat**—in this kind of distractor, heat is characterized by the attributes of materialistic substances, according to Reiner *et al.* and Eshach [10]. For instance, distractor (c) in item 1 suggests a "torch emits heat particles that fly in all directions. When these particles hit our hand, they heat it up." Here, heat is materialistic (particles) and transitional—flows from A to B. Distractor (a) in item 7 claims that we feel hot on a sunny day since, "heat emitted from the sun accumulates in the atmosphere, which, as a result, heats up and warms the earth." In this explanation, heat is containable, that is, can be accumulated in the atmosphere.
2. **Hot atoms or molecules**—in this kind of distractor, hot atoms or molecules travel from one location to another [5,20]. For instance, distractor (d) in item 1 suggests regarding heat conduction in the air that, "Fire heats the air molecules nearby and causes them to oscillate strongly. These oscillating molecules fly in all directions. Some of them hit the hand and irritate the skin."

TABLE II. Relating the NOHI possible answers to categories.

Item	<i>Materialistic heat</i>	<i>Hot atoms or molecules</i>	<i>Expanding atoms or molecules</i>	<i>Kinetic molecular theory</i>
1	c	d	a	b
2	a	c	d	b
3	c	d	a	b
4	a	c	d	b
5	a	b	c	d
6	b, e	d	c	a
7	a	d	c	b
8	c	a	b	d
9	b	a	d	c
10	b	d	a	c

3. **Expanding atoms or molecules**—these distractors are based on the idea that atoms or molecules of a heated-up material expand (like a balloon); and they shrink in a cooled one. For instance, distractor (b) in item (8) suggests that the liquid column in the tube of a thermometer gets higher, since the “liquid’s molecules warm up and therefore their volume increase.”

All of the distractors were constructed based on these three categories. The ideas and misconceptions identified in both focus group and clinical interviews were actually used to formulate about 80% of the distractors. The other 20% were based on known misconceptions from the history of science and the authors’ field experience as physics lecturers. Responses based on relatively advanced and/or abstract terms, such as energy, entropy, Casimir effect, etc., were omitted from the distractors. The correct answers are associated with the kinetic molecular theory of matter and electromagnetic radiation when necessary in layman’s terms. We relate all the possible responses in NOHI to these categories in Table II below. As one can see, all the items include one response of each category.

V. FIELD TESTING—RELIABILITY ESTIMATION, ITEM ANALYSIS

We used the classical test theory to analyze the following three psychometric features of NOHI: Cronbach’s α reliability, item discrimination index, and item point biserial coefficient [13,17,29]. *Test reliability* is an indicator of how precisely the test makes the measurement [13]. A common way of measuring the internal consistency of a test is by calculating the Cronbach’s alpha coefficient ($0 \leq \alpha \leq 1$). Conceptual knowledge tests having $\alpha \geq 0.7$ are generally considered to be reliable [29]. *Index of discrimination* (D) indicates how well an item distinguishes top-scoring students from poorly performing ones [6,29]. Janda [29] located the minimal D level at 0.28. *Point biserial coefficient* (r_{bis}) indicates how consistent each item is with the whole assessment. High point biserial

coefficient values usually provide additional evidence of the item’s quality [13,29]. The minimal acceptable level of r_{bis} is 0.2 [6,32].

The first draft of NOHI was administered to 59 participants: undergraduate engineering students, B.Sc., M.Sc., and Ph.D. teachers for natural sciences and engineers. According to the literature, a minimal sample of 20 or 30 people is usually enough to get some useful statistics. The numbers start becoming statistically stable near 50 examinees [33–35]. Therefore, a starting sample of 59 participants was considered sufficient. In light of the item analysis carried out on the first draft, “unpopular” distractors that were chosen by less than 10% of the participants were omitted or rephrased. For instance, we considered that the HOS idea of heat as an invisible liquid (caloric) was chosen by less than 4%. Therefore, we decided to remove all the distractors relating to this idea and include it in the *materialistic heat* category. At the same time, we made sure that all the items included all the four categories of answers so the instrument would be suitable for use in the terms of mental model analysis of Bao [24,36].

In order to estimate the reliability of the second version of NOHI and provide further item fine-tuning, we distributed the instrument to 81 participants. Of the participants, 47 studied on B. Ed. Program for natural sciences teachers and took a basic physics course including mechanics and classical heat theory prior to answering NOHI; 21—engineering students; and 13 B.Sc., M.Sc., and Ph.D. graduated physicists and engineers. This group yielded an average grade of 36% (SD = 26.9%) and Cronbach’s $\alpha = 0.69$. Table III below presents the index of discrimination and point-biserial coefficient for each item. As can be seen in the table, the mean values of both meet the literature requirements as previously presented.

Figure 1 depicts the distribution of the participants among the categories of distractors. The histogram

TABLE III. Index of discrimination and point biserial coefficient calculated for NOHI items.

Item	D	r_{bis}
1	0.69	0.62
2	0.53	0.66
3	0.63	0.65
4	0.63	0.6
5	0.28	0.38
6	0.31	0.39
7	0.53	0.54
8	0.63	0.64
9	0.63	0.57
10	0.56	0.85
Mean	0.54	0.56
SD	0.14	0.1

actually represents the average probability of a participant in choosing an answer of a certain category in an item.

VI. ESTABLISHING VALIDITY

The issue of test validity addresses the questions: Does the test measure what we think it does? To what extent does it succeed in measuring [29]? The *content-related aspect* of validity is probably the most important for knowledge examining tests like the one we developed in the present study. It deals with the question: To what extent does the content of the test match the domain of interest. Thus, a test is considered content-valid if its content corresponds to a student's performance that we want to observe [6,10]. We estimate content validity in the following two ways:

1. *Specifications table*—One of the methods for establishing content validity of a test is to outline test specifications, which denote the content domains that the test intends to encompass. One way to relate test specifications is to construct a two-dimensional test specifications table, that is, topics in the test vs the appropriate items [6,29]. Table I above is actually such a specification table. In this table, we showed that NOHI relates to most of the topics in the nature of heat described by the kinetic molecular model as conveyed in undergraduate physics textbooks, such as Halliday *et al.* [37].
2. *Experts' review*—deals with experts' review of the content [13,29]. The professional literature provides a rough consensus of the number of experts, 3–10 being sufficient for estimating content validity. Indeed, Rubio *et al.* [38] recommended including 3–10 experts, Volfson *et al.* [6] used 6, and Eshach [10] used 7. Twelve experts reviewed NOHI as specified in phase 1. We asked the experts to (a) complete the test, (b) comment on the appropriateness of the objectives, and (c) criticize the items. They were also invited to comment on the format and on how the phrasing of the statements could be improved. The experts and authors then discussed the comments and suggestions. As a result, we modified or removed some items/distractors according to the experts' recommendations.

Face validity exists if a test looks as if it relates to what it has been declared to relate to and measure from the examinees' point of view [29]. Indeed, poor face validity may negatively affect the attitude and motivation of participants to cooperate with the examiner. To provide face validity of an instrument, an examiner usually gives the test for appreciation to a sampling group of the future target audience [6,10]. To establish face validity, we discussed the instrument with five practicing physics lecturers. They approved the instrument's face validity, and we considered their opinion as being satisfactory.

VII. DISCUSSION

In the present study, we developed the nature of heat inventory measuring students' understanding of basic heat phenomena in terms of classical kinetic molecular theory. To the best of our knowledge, such diagnostic instruments are absent in the professional literature. NOHI aims to fill this gap. Following Adams and Weiman [13] as well as previous studies [6,10], the development of NOHI included formulating the objectives, constructing the items, and performing reliability and validity checks. A variety of sources were used to design the distractors: literature review, HOS, responses by undergraduate and graduate students in personal and focus-group interviews, participants' responses to the first and second drafts of NOHI, physics textbooks, and a group of 12 experts, including experienced physics teachers and lecturers having an extensive physics background. The experts were asked not only to examine the validity of NOHI but also to suggest other items and/or distractors and rephrase the existing ones. This integration of multiple sources is in line with Creswell [39], who suggests that we must “discuss plans to triangulate, or find convergence among sources of information, different investigators, or different methods of data collection” (p. 158) in order to increase the internal validity of an instrument. The validity was found to be sufficient.

A. Statistical parameters of NOHI

The second version of NOHI yielded $\alpha = 0.69$, which is almost the lower acceptable limit according to Janda [29]. Moreover, many researchers believe that a high Cronbach's α obtained in a knowledge test such as NOHI does not indeed guarantee high reliability but may actually be an indication that there are redundant items that should be removed, whereas a relatively low alpha coefficient would be quite reasonable [13,40,41]. For instance, Volfson *et al.* [6] reported $\alpha = 0.66$. Berger and Hänze [40] obtained $\alpha = 0.6$, which they considered to be acceptable

“...given the limited number of test items and the broad range of tested knowledge” (p. 303).

The same can be expected for NOHI, which also measures a range of concepts (see Table I). Thus, we consider $\alpha = 0.69$ as sufficient.

According to Janda [29], the acceptable lower border of *index of discrimination* is 0.28. All of the items in NOHI produced $D \geq 0.28$, and the average index of discrimination was 0.54 (see, Table III), which meets the literature recommendations. The *point biserial coefficients* of the items varied between 0.39 and 0.66, indicating a degree of correlation of each item with the whole assessment. These results meet the recommendation of Adams and Weiman [13], arguing that the *point biserial coefficient* should be positive but not too high.

NOHI yielded an average grade of 36% during its field testing. This might appear relatively low. However, NOHI does not just require an understanding of science concepts, but also demands their application for the analysis of practical situations. The fact that *application* and *analysis* are both placed above *understanding* in Bloom's taxonomy [42] automatically makes NOHI more difficult than other instruments intended for the same level of knowledge. Indeed, it has been found that people have difficulty applying abstract scientific concepts and rules to practical situations [43]. Thus, the relatively low average grade appears to be quite reasonable and can indicate a lack of experience in practical application of the learnt science material. On the other side, it might also indicate the need in more extensive instruction of the molecular kinetic and classical heat theory already on the middle-school level, even more so on the academic level especially, for future science teachers.

B. Categories discussion

Nevertheless, we ran NOHI in groups of future sciences teachers, who definitely learned classical heat theory, at least qualitatively; the most popular category of misconceptions was still the *materialistic* one (24.3%). The hybrid (materialistic + particles' motion) *hot atoms* category yielded 18.3%. That is in total 42.6% probability of a participant to be triggered into materialistic way of thinking regarding heat processes by one of NOHI items. This finding indicates *materialistic* misconceptions to be robust and resistant for teaching, thus reinforcing the findings of Chi [2]. Consequently, instructors should make special efforts to address materialistic thinking and facilitate reviewing such misconceptions, that is, by contradictory demonstrations, challenging questions and explanation of the inner mechanism of emergent processes such as the heat ones [3,5].

As to the third category of misconceptions—*expanding atoms*, this mental model can be explained by the everyday physical experience of the participants. Indeed, all of usual materials expand when heated (except of the water anomaly). This is a “piece of knowledge,” a *p-prim* (phenomenological primitive) in the terms of DiSessa. P-prims are elements of knowledge that constitute “people's “sense of mechanism,” their sense of which happenings are obvious, which are plausible, which are implausible, and how one can explain or refute real or imagined possibilities” [44]. P-prims are usually activated when interpreting scientific phenomena. Herein, naïve students mistakenly relate the p-prim “*heating causes expansion*” to the microscopic—atomic level of matter. Nevertheless, according to DiSessa, naïve knowledge is noncoherent and the set of p-prims strongly varies from person to person, the *expanding atoms* model is statistically significant. This can be explained by the fact that thermal expansion is a very popular everyday physical experience in the modern society, for instance, we see it

when boiling water or check the oil level in a hot vs cold engine of a car. These experiences form the same p-prim in the minds of most of the population [19].

C. Methodological limitations

First, it should be emphasized that NOHI was developed basing on the Israeli population of physicists, engineers, and students. The results might be different if we had recruited participants from other countries. Indeed, participants from other countries might be exposed to physics education at a different level and scope of topics, different demonstrations and experiences. On the other hand, NOHI can be used to examine such differences [10].

Second, NOHI is intended to measure understanding of the very basic issues in classical heat theory. It does not relate on purpose the more advanced and still very important thermodynamic concepts such as specific heat, temperature, energy, entropy, etc. It is the time now to develop next inventories, which will address them.

VIII. CONCLUSION

We started this paper emphasizing the importance of kinetic molecular theory in science instruction and in contrast, the widespread difficulties of students in its practical implementation. The average grade of NOHI (36%) in future science teachers' sample reinforces the strong and immediate need to develop practical pedagogic instruments for more efficient kinetic molecular theory teaching. We hope NOHI to meet this need at least in classical heat theory domain. Actually, NOHI can contribute to three aspects in the field of science education: diagnostic, instruction, and research.

Actually, diagnostics of thinking is the basic purpose of conceptual inventories such as NOHI. It can be conducted within a group of learners before almost every thermodynamics course to identify the most popular misconceptions, which further have to be treated by the instructor. It also can be conducted while a course as formative assessment and at the end of a course as a summative assessment.

In the instruction field, NOHI provides three practical resources:

- The items' distractors actually provide a detailed list of most frequent misconceptions regarding heat processes. There is general agreement today among science educators that teachers' deep understanding of their students' knowledge is a necessary key step to (a) design effective learning environments and; (b) lead learners to conceptual change from the non-correct mental models to the correct ones [6,7,10,19]. We hope, therefore, science educators will find benefit in NOHI's distractors.
- An ensemble of inspiring and challenging questions for whole class dialogic discussions [43], problem-based learning assignments [45,46], etc.

- c. Ideas for demonstration experiments in physics classes. For instance, see items 4, 5, 6, and 9 (Appendix).

As to research point of view, NOHI can be used as a large-scale assessment tool, the results of which may lead to discourse about teaching methods and topics that should be included in the curriculum to meet the requirements of modern education. NOHI may also help to address some research questions, such as whether differences in understanding of heat processes exist among students studying in different faculties: physics, chemistry, biology, science teachers, etc. It could also help to assess novel learning environments in preinstruction and postinstruction analysis [47]. Finally, we hope that NOHI shall provide an inspiration for developing similar inventories intended to measure students' ideas regarding the more advanced thermodynamic concepts.

DATA AVAILABILITY

The data that support the findings of this article are openly available [22].

APPENDIX A: NATURE OF HEAT INVENTORY

1. **Place your hand next to a burning torch at some distance from it. The hand does not touch the flame, but you feel hot. Why?**



- Air molecules around the torch heat up and expand. Because of the increase in volume, they hit neighboring molecules. The neighbors also expand, and thus, heat reaches the hand.
- Heat is transferred through accelerated movement and increased collisions of air molecules, as well as by electromagnetic radiation that moves the molecules and electrons in the hand's skin.
- The torch emits heat particles that fly in all directions. When these particles hit the hand, they heat it up.
- Fire heats the air molecules nearby and causes them to oscillate strongly. These oscillating molecules fly in all directions. Some of them hit the hand and irritate the skin.

- e. None of the above choices fits my basic viewpoint. My basic viewpoint is _____.

2. **Place your hand above a torch. Now it feels much hotter. If you hold your hand over the flame for a few seconds, you will get a burn. Why does heat go up?**



- Heat is lighter than air and therefore rises up.
 - Heating causes increased collisions between air molecules. As a result, the density of the hot air decreases and it flows up.
 - Hot air molecules are lighter than cold ones and therefore go up.
 - Hot air molecules expand. Thus, their density decreases and they flow up like air bubbles in water.
 - None of the above choices fits my basic viewpoint. My basic viewpoint is _____.
3. **A person holds an ice cube in his hand. He feels cold in his hand and the cube melts. Why?**
- The volume of cold ice molecules is small. As soon as they contact the warm hand's molecules, ice molecules expand and turn into water. The molecules of the hand, on the other hand, shrink.
 - Molecules in the tissues of the hand vibrate stronger than the ice molecules. Skin molecules collide with ice molecules, oscillate them and thus, ice melts.
 - Heat flows from the hand to ice, enters between the ice molecules and thus increases its volume. When heat fills enough spaces between the ice molecules, the bonds between them break and solid becomes a liquid.
 - Ice molecules in the outer layer heat up and accelerate, into the cube. They enter inside, breaking the bonds between the cold molecules.

- e. None of the above choices fits my basic viewpoint. My basic viewpoint is _____.

4. **We hang a copper rod above a torch and allow it to heat up. The rod lengthens by 2 mm. Why?**



- a. Heat emitted from the torch enters between the copper atoms, fills the spaces between them and thus increases the volume of the rod.
 b. Heating causes increased movement and collisions of the atoms and electrons in the rod, and therefore, it lengthens.
 c. The copper atoms in the outer layer heat up and accelerate into the rod. They collide with the cold atoms inside, causing them to move sideways.
 d. The copper atoms expand as a result of the heating, and therefore, the length of the rod increases.
 e. None of the above choices fits my basic viewpoint. My basic viewpoint is _____.

5. **We place a candle close to a torch. What will happen to the flames of both?**



- a. The torch emits more heat, the candle less. Therefore, the candle flame will deviate strongly to one side. The torch flame will deviate slightly to the other side.
 b. The larger torch heats and accelerates more air molecules around it. These molecules will hit

the candle flame and cause it to deviate strongly to the side.

- c. Air molecules around both flames will expand due to the heating. This will cause both flames to deviate to the sides.
 d. Heat of the flames cause the air molecules' collisions to be stronger and more frequent. However, since these collisions occur randomly in all directions, they will not influence the flames' shape.
 e. None of the above choices fits my basic viewpoint. My basic viewpoint is _____.

6. **We stick two copper rods: a thin one and a thick one, of equal length, into the body of an unlit candle (made of paraffin). Now we heat the ends of the rods found in the air. Heat will pass through each of the rods, melting the paraffin around its other end, and the rod will fall. Which rod will fall first?**

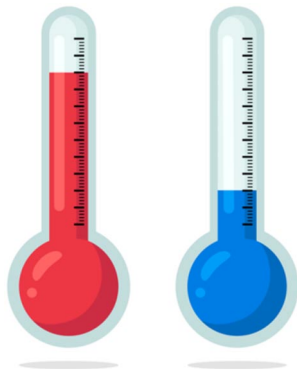


- a. Heating will cause increased movement and collisions between the copper electrons in random directions. Therefore, the thicker rod will transfer heat faster and fall first.
 b. Because of heating, the copper atoms expand one after the other. The rate of expansion is constant in the same material, and therefore, the rods will fall together.
 c. Copper atoms accelerated by heating will move to the other end at the same speed in both rods. Therefore, the rods will fall together.
 d. The volume of the thin rod is smaller, therefore, the heat will fill it faster and it will fall first.
 e. None of the above choices fits my basic viewpoint. My basic viewpoint is _____.

7. **On a sunny day, we feel hot. How does the sun heat the Earth's surface?**

- Heat emitted by the Sun accumulates in the atmosphere, which, as a result, heats and warms the Earth.
- The Sun creates time-varying electromagnetic radiation, which oscillates the molecules and electrons in the materials in the ground.
- Sunlight causes air molecules in the atmosphere to expand. As a result, they collide with each other more, thus transferring heat to the ground.
- Solar radiation hits the air molecules in the atmosphere and accelerates them towards the Earth's surface. They collide with the ground and heat it.
- None of the above choices fits my basic viewpoint. My basic viewpoint is _____.

8. **A thermometer is an instrument for measuring temperature. A simple thermometer is made of a thin glass tube containing a liquid (mercury, alcohol, etc.). The higher the body temperature, the higher is the column of liquid in the tube. Why?**



- The liquid molecules near the tube walls heat up. They are lighter than the rest of the liquid and therefore rise to the top. As a result, the height of the liquid grows up.
- The liquid molecules heat up and, as a result, expand. Therefore, the volume of the liquid increases.
- Heat generated by the body enters between the liquid molecules and thus increases its volume.
- Body heat causes stronger movement and collisions of the liquid molecules. As a result, the volume of the liquid increases.
- None of the above choices fits my basic viewpoint. My basic viewpoint is _____.

9. **We hang a weight on a rubber band. The rubber stretches. We heat the air around the rubber. The rubber band will get shorter and pull the weight up. Why?**



- Hot air molecules move faster than the rest, collide the rubber strongly and cause the rubber to zig-zag, which reduces its length.
- Because of the accumulation of heat in the rubber, a chemical reaction occurs between the rubber molecules, and therefore, the distance between the molecules decreases and they "stick" to each other.
- A rubber band is similar in structure to a chain. Heating causes increased movement of the rubber band molecules to the right and left, which reduces its length.
- The rubber band molecules expand due to heating, entering into a chemical interaction with each other, which reduces their volume.
- None of the above choices fits my basic viewpoint. My basic viewpoint is _____.

10. **We melt iron into a liquid state and then allow it to cool. The iron will become solid. Its volume will decrease. How?**

- The iron atoms' volume decrease because of cooling and they "stick" to each other.
- Heat is released from the interatomic spaces. As a result, the iron cools and the atoms come closer each other.
- While cooling, conduction electrons' velocities decrease and their collisions with the surrounding atoms become weaker thus, causing the interatomic distances to decrease.
- Hot conduction electrons are emitted from the iron. As a result, hot iron radiates red light and its volume decreases as it cools.
- None of the above choices fits my basic viewpoint. My basic viewpoint is _____.

- [1] R. P. Feynman, *The Feynman Lectures on Physics* (Addison-Wesley, Reading, MA, 1963), Vol. I, Ch. 1.
- [2] M. T. H. Chi, Three types of conceptual change: Belief revision, mental model transformation, and categorical shift, in *Handbook of Research on Conceptual Change*, edited by S. Vosniadou (Erlbaum, Hillsdale, NJ, 2008), pp. 61–82.
- [3] M. T. H. Chi, R. D. Roscoe, J. D. Slotta, M. Roy, and C. C. Chase, Misconceived causal explanations for emergent processes, *Cogn. Sci.* **36**, 1 (2012).
- [4] D. A. Nugraha, A. Suparmi, R. Winarni, and Suciati, Students' profile of heat and temperature using HTCE in undergraduate physics, *J. Phys.: Conf. Ser.* **909**, 012059 (2017).
- [5] A. Volfson, H. Eshach, and Y. Ben-Abu, Introducing the idea of entropy to the ontological category shift theory for conceptual change: The case of heat and sound, *Phys. Rev. Phys. Educ. Res.* **15**, 010132 (2019).
- [6] A. Volfson, H. Eshach, and Y. Ben-Abu, Development of a diagnostic tool aimed at pinpointing undergraduate students' knowledge about sound and its implementation in simple acoustic apparatuses analysis, *Phys. Rev. Phys. Educ. Res.* **14**, 020127 (2018).
- [7] S. Vosniadou, *International Handbook of Research on Conceptual Change*, 2nd ed. (Routledge/Taylor & Francis Group, New York, 2013).
- [8] K. E. Dunn and S. W. Mulvenon, A critical review of research on formative assessments: The limited scientific evidence of the impact of formative assessments in education, *Pract. Assess. Res. Eval.* **14**, 1 (2009).
- [9] P. Johnston and P. Costello, Principles for literacy assessment, *Read. Res. Q.* **40**, 256 (2005).
- [10] H. Eshach, Development of a student-centered instrument to assess middle school students' conceptual understanding of sound, *Phys. Rev. ST Phys. Educ. Res.* **10**, 010102 (2014).
- [11] H. Eshach and I. Kukliansky, Developing of an instrument for assessing students' data analysis skills in the undergraduate physics laboratory, *Can. J. Phys.* **94**, 1 (2016).
- [12] D. Hestenes, M. Wells, and M. Swackhamer, Force concept inventory, *Phys. Teach.* **30**, 141 (1992).
- [13] W. K. Adams and C. E. Wieman, Development and validation of instruments to measure learning of expert-like thinking, *Int. J. Sci. Educ.* **32**, 1 (2010).
- [14] B. Orazov, G. Issayeva, S. Maxutov, E. Kozhabekova, and N. Balta, Exploring university students' misconceptions of the kinetic molecular theory of gases: A study from Kazakhstan, *Eurasia J. Math. Sci. Technol. Educ.* **21**, em2587 (2025).
- [15] S. R. Pathare and H. C. Pradhan, Students' misconceptions about heat transfer mechanisms and elementary kinetic theory, *Phys. Educ.* **45**, 629 (2010).
- [16] R. K. Thorton and D. K. Sokoloff, Heat and Temperature Conceptual Evaluation, [physport.org/assessments/assessment.cfm?I=16&A=HTCE](https://www.physport.org/assessments/assessment.cfm?I=16&A=HTCE) (2001).
- [17] L. Ding, R. Chabay, and B. Sherwood, How do students in an innovative principle-based mechanics course understand energy concepts?, *J. Res. Sci. Teach.* **50**, 722 (2013).
- [18] S. Yeo and M. Zadnik, Introductory thermal concept evaluation: Assessing students' understanding, *Phys. Teach.* **39**, 2001.
- [19] A. Volfson, H. Eshach, and Y. Ben-Abu, Physics teaching via dialogic discussions about circus activities, *Contemp. Educ. Technol.* **15**, ep421 (2023).
- [20] M. T. H. Chi, Commonsense conceptions of emergent processes: Why some misconceptions are robust, *J. Learn. Sci.* **14**, 161 (2005).
- [21] S. Vosniadou and W. F. Brewer, Mental models of the day/night cycle, *Cogn. Sci.* **18**, 123 (1994).
- [22] D. L. Medin and L. J. Rips, Concepts and categories: Memory, meaning, and metaphysics, in *The Cambridge Handbook of Thinking and Reasoning*, edited by K. Holyoak and B. Morrison (Cambridge University Press, Cambridge, England, 2005), pp. 37–72.
- [23] A. A. DiSessa, Knowledge in pieces, in *Constructivism in the Computer Age*, edited by G. Forman and P. Pufall (Lawrence Erlbaum, New Jersey, 1988), pp. 49–70.
- [24] L. Bao, Dynamics of student modeling: a theory, algorithms and application to quantum mechanics, University of Maryland, College Park, 1999.
- [25] M. Reiner, J. D. Slotta, M. T. H. Chi, and L. B. Resnick, Naïve physics reasoning: A commitment to substance-based conceptions, *Cogn. Instr.* **18**, 1 (2000).
- [26] E. Hepler-Smith, *The Elements: A Visual History of Their Discovery* (review of Book by Philip Ball (University of Chicago Press, Chicago, 2022)). ISBN: 978-0-226-77595-1.
- [27] H. Eshach and J. L. Schwartz, Sound stuff? Naive materialism in middle-school students' conceptions of sound, *Int. J. Sci. Educ.* **28**, 733 (2006).
- [28] A. L. Chandrasegaran, D. F. Treagust, and M. Mocerino, The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation, *Chem. Educ. Res. Pract.* **8**, 293 (2007).
- [29] L. H. Janda, *Psychological Testing: Theory and Applications* (The Open University, Tel Aviv, 2008).
- [30] R. M. Dilshad and M. I. Latif, Focus group interview as a tool for qualitative research: An analysis, *Pak. J. Soc. Sci.* **33**, 191 (2013).
- [31] N. Dogan and F. Abd-El-Khalick, Turkish grade 10 students' and science teachers' conceptions of the nature of science: A national study, *J. Res. Sci. Teach.* **45**, 1083 (2008).
- [32] R. Beichner, Testing student interpretation of kinematics graphs, *Am. J. Phys.* **62**, 750 (1994).
- [33] J. C. Cappelleri, J. J. Lundy, and R. D. Hays, Overview of classical test theory and item response theory for quantitative assessment of items in developing patient-reported outcome measures, *Clin. Ther.* **36**, 648 (2014).
- [34] L. Terry and K. Kelley, Sample size planning for composite reliability coefficients: Accuracy in parameter estimation via narrow confidence intervals, *Br. J. Math. Stat. Psychol.* **65**, 371 (2011).

- [35] N. A. Thompson, *Introduction to Classical Test Theory with CITAS* (Assessment Systems Corporation, Minnetonka, 2016).
- [36] L. Bao and E. F. Reddish, Model analysis: Representing and assessing the dynamics of student learning, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010103 (2006).
- [37] D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*, 10th ed. (Wiley, New York, 2014).
- [38] D. M. Rubio, M. Berg-Weger, S. S. Tebri, E. S. Lee, and S. Rauch, Objectifying content validity: Conducting a content validity study in social work research, *Soc. Work Res.* **27**, 94 (2003).
- [39] J. W. Creswell, *Research Design: Qualitative and Quantitative Approaches* (SAGE Publications, Thousand Oaks, CA, 1994).
- [40] R. Berger and M. Hänze, Impact of expert teaching quality on novice academic performance in the jigsaw cooperative learning method, *Int. J. Sci. Educ.* **37**, 294 (2015).
- [41] K. S. Taber, The use of Cronbach's alpha when developing and reporting research instruments in science education, *Res. Sci. Educ.* **48**, 1 (2016).
- [42] A. Batey, J. Pollard, S. Shott, J. Yost, and K. Merrill, Development of learners' thinking by means of information technologies, in *Intel Program "Future learning"*, edited by E. N. Yastrebtzev (Printing Plant of Bogorodsk, Noginsk, 2006).
- [43] H. Eshach, An analysis of conceptual flow patterns and structures in the physics classroom, *Int. J. Sci. Educ.* **32**, 451 (2009).
- [44] A. A. DiSessa, A friendly introduction to "Knowledge in pieces": modeling types of knowledge and their roles in learning, in *Invited Lectures from the 13th International Congress on Mathematical Education*, edited by H. Kaiser *et al.* (Springer, Switzerland, 2018), pp. 65–84.
- [45] D. Raine and J. Collett, Problem-based learning in astrophysics, *Eur. J. Phys.* **24**, 41 (2003).
- [46] A. Volfson and Y. Ben-Abu, The physics of a rope dart: Mass, force and circular motion, *Phys. Teach* **62**, 2024 (2024).
- [47] N. S. Rebello and D. A. Zollman, The effect of distracters on student performance on the force concept inventory, *Am. J. Phys.* **72**, 116 (2004).