Probability-related naïve ideas across physics topics

Michael M. Hull, Alexandra Jansky & Martin Hopf

Published online: 12 May 2020.

Submit your article to this journal ⍰

Article views: 929

View related articles ⍰

View Crossmark data ⍰

Citing articles: 1 View citing articles ⍰

Routledge
Taylor & Francis Group

Probability-related naïve ideas across physics topics

Michael M. Hull [a], Alexandra Jansky [a,b] and Martin Hopf[a]

[a]Austrian Educational Competence Centre Physics, University of Vienna, Vienna, Austria; [b]CERN, European Organization for Nuclear Research, Geneva, Switzerland

ABSTRACT

In this literature review, we survey student naïve ideas (frequently referred to as 'misconceptions') that plausibly relate, at least in part, to difficulty in understanding probability. We collected diverse naïve ideas from a range of topics in physics: Non-linear Dynamics; Cosmology; Thermal Physics; Atomic, Nuclear, and Particle Physics; Elementary Particle Physics; Quantum Physics; and Measurements and Uncertainties. With rare exception, these naïve ideas are treated in the literature to be topic-specific. For example, the idea that 'only one measurement is needed because successive measurements will always yield the same result' is treated to be a misconception in Measurements and Uncertainties. In our review, however, we raise the possibility that these diverse naïve ideas have something in common: they are enabled, to varying degrees, by the stance that 'random is incompatible with predictions and laws' that researchers in mathematics education have documented. This is important, as it may inform instruction. Namely, it may be the case that it is more effective to treat this underlying cause of student difficulty, rather than the individual naïve ideas themselves.

Introduction

To date, the field of physics education research has published findings of the existence and resilience of thousands of students' naïve ideas spanning virtually every topic of physics. (Duit, 2009; Flores et al., 2014; McDermott & Redish, 1999) Naïve ideas tend to be described within a single physics topic, without discussion of similar ideas in other topics. This is likely due in part to how physics courses themselves are structured. Due to the breadth of material covered by the field of physics and regarded as important by physicists and physics teachers alike, many students take several physics classes, with content separated between the classes. For example, at the college level, mechanics is typically taught in a different course than electromagnetism. Within a given course, content is further segmented into topics such as kinematics, dynamics, energy, momentum, torque, etc. In this literature review, we present publications on naïve ideas from diverse physics topics that seem to share a common feature: they plausibly stem, at least in part, from student difficulty in under-standing the concept of probability. The notion that these naïve ideas might have a common origin is relevant for educators. Namely, if it turns out that this is indeed the case, it would imply that effective instruction for students with these naïve ideas would

---

involve focusing on the underlying cause, instead of trying to treat many symptoms as though they are unrelated.

Literature, particularly from mathematics education, has documented the existence of myriad student difficulties in understanding probability and randomness (*e.g.* Barragués et al., 2006; Batanero et al., 2016; Buechter et al., 2005; Döhrmann, 2005; Green, 1983; Kahneman & Tversky, 1973; Shaughnessy & Ciancetta, 2002; Shaughnessy & Zawojewski, 1999; D. Stavrou et al., 2003). Perhaps because of the emphasis given in the classroom on coins, dice, and cards for teaching probability, students often assume that something must have equal probability of occurring (1 out of 52 for drawing a particular card, for example) in order for it to qualify as 'random'. (Buechter et al., 2005) A well-known naïve idea is gambler's fallacy, the idea that, for example, if a roulette wheel has been black the past several times in a row, then it is more likely to be red with the next spin. (Buechter et al., 2005; Green, 1983) The opposite naïve idea has also been documented, that when a good luck streak is broken, people search for an explanation (instead of understanding that it is regression to the mean). (Kahneman & Tversky, 1973) These two naïve ideas suggest a hesitancy for people to accept the idea that things can actually be random. In the case of the roulette wheel, for example, it is as though there is a belief in an invisible but intentional agent at work ensuring that the next spin will be red to 'compensate' for all the black prior results. Consistent with this is the idea that things only *appear* random because there is insufficient information to know for sure what will happen. (Barragués et al., 2006; Döhrmann, 2005; D. Stavrou et al., 2003) Plausibly indicating this view, when asked what would be necessary to be able to predict where a ball travelling down a plinko board (a vertically hanging board with many pegs protruding from it) would land at the bottom, few students described the approach of repeating the experiment many times to see which outcome is the most likely. Rather, students attempted to remove the aspect of randomness by saying that one should get more details, for example, about the material that the ball and pegs are made out of.[1] (Barragués et al., 2006) The idea that a prediction cannot be made from situations where randomness is involved has been described for other situations as well (e.g. Stavrou & Duit, 2014).

It seems that for many students, 'random' is not compatible with 'lawful/predictable' (Buechter et al., 2005; Gougis et al., 2017; Stavrou et al., 2003) in the sense that things can *either* be predicted in their entirety with the governing laws *or* that there are no rules and nothing is certain, making predictions useless. However, much of physics concerns making predictions about phenomena that are intrinsically uncertain but nevertheless are governed by laws. In a sealed gas, for example, the force an individual molecule exerts on a wall at a given moment is considered random. Nevertheless, we predict that the pressure of the gas will increase as the container is compressed, in accordance with the ideal gas law. This literature review concerns itself with physics topics in which ideas about randomness and probability are relevant. We find it plausible that many of the seemingly disparate naïve ideas discussed in this paper are related to each other, in that they stem from this stance of 'random is incompatible with predictable'. Despite this common feature, of the 70 articles reviewed (15 in Non-linear Systems, 2 in Cosmology, 34 in Measurements and Uncertainties, 9 in Thermal Physics, and 10 from the 'Structure of Matter' heading) only 9 of them mentioned student difficulties in multiple fields, and then only briefly.

Many of the naïve ideas that we review here (for example, that students often conflate heat and temperature) are likely to be familiar to the reader, and so we will not provide unnecessarily detailed descriptions of the ideas. Of course, we provide references to the works we cite so that interested readers can find more description of the naïve ideas and the studies that documented them. We will, however, interpret these ideas, from the perspective that they may in part be the result of the stance that 'random is incompatible with predictable'. To assist our interpretation, we will first briefly discuss the ontologies framework, which will give credence to our hypothesis that the naïve ideas we review, despite being in disparate physics topics, could be due to a common cause.

The ontologies framework

In describing misconceptions[2] and conceptual change, Chi discussed how misconceptions can vary in how resilient they are. A 'false belief', for example, that the heart oxygenates blood, can often be easily corrected just by telling students that the lungs oxygenate blood. Other times, however, students will have 'an organized collection of individual beliefs' that, despite being normatively incorrect, are nevertheless self-consistent with each other. (Chi, 2013) At times, these collections of beliefs can be expansive and self-coherent to the extent that they can be referred to as 'theories'. (e.g. Carey, 2009; Lattery, 2016; Strike & Posner, 1982; Vosniadou & Skopeliti, 2014) For example, students learning mechanics are often consistent in their views that a moving object carries a force propelling it in the direction of motion, which is comparable in some ways to the 'impetus theory' of Galileo. (e.g. Lattery, 2016) Chi argues that the misconceptions that are most resistant to correction are those resulting from learners assigning the relevant physics concept into the wrong *ontology*, or mental category into which an individual divides the world. (Keil, 1979) Two examples of ontologies are 'entities' (with which people categorise things like sharks and water) and 'process' (like a baseball game or evolution). A 'shark' is not only a distinctly different thing than a 'baseball game'; it is a distinctly different *type* of thing. Evidence for this comes from the different predicates used to describe entities in the two ontologies. For example, whereas a 'process' has the attribute of 'having certain duration', 'entities' do not: although a 'baseball game' can be 'two hours long', it is nonsensical to talk about a 'shark' in this way. Chi et al. argue that many concepts such as electric current, heat, and light are difficult for students to learn because, whereas the scientific understanding assigns these concepts into the ontology of 'process', learners tend instead to see them as 'entities' (termed 'substances' or 'matter' in earlier work). (Chi, 2013; Chi & Slotta, 1993; Chi et al., 1994; Reiner et al., 2000; Slotta & Chi, 2006; Slotta et al., 1995)(Chi et al., 1994; Chi & Slotta, 1993; Reiner et al., 2000; Slotta & Chi, 2006; Slotta et al., 1995)

Within each ontology is nested sub-ontologies. Although Chi et al. specify that they are 'not committed to this exact hierarchy' (Chi & Slotta, 1993), they provide an example hierarchy in which 'event' is nested within 'process' and 'random' nested as a type of 'event'. In light of the findings discussed above that learners find 'random' things to be incompatible with predictions, we find it plausible that an additional sub-ontology with which people categorise 'events' is one that attributes lawfulness and predictability to the event. We will call this hypothetical ontology the 'deterministic' ontology.

The 'deterministic' ontology is a mental category people use to classify events they perceive to be governed by rules, demonstrate a pattern, and lead to predictable outcomes. Events classified in this ontology do *not* exhibit randomness. The motion of a projectile in an introductory physics class can be assigned into this ontology, and it is thus appropriate to use kinematic equations to predict the trajectory.

The 'random' ontology is a mental category people use to classify events they perceive to be lawless, patternless, and with unpredictable outcomes. For these events, equations and other models are *not* useful.[3] Chi and Slotta give the example of 'mutation' for this ontology. (Chi & Slotta, 1993) The time at which a radioactive nucleus decays can also be assigned into this ontology.

The ontologies framework seems promising to characterise the wide spread of naïve ideas reviewed in this paper for two reasons. First, whereas a naïve idea or theory is specific to a given physics topic (for example, a naïve idea about 'force' is specific to the topic of mechanics), an ontology spans across topics, which is a necessary criteria for describing the underlying probability-related difficulty we hypothesise to be behind the naïve ideas we review. That is, Chi et al. argued that disparate naïve ideas relating to electricity, thermodynamics, and light have a shared cause of learners assigning the phenomena incorrectly into the 'entities' ontology. We likewise argue that a variety of probability-related naïve ideas may have a shared cause of learners assigning the phenomena incorrectly into either the 'random' or 'deterministic' ontologies. The second reason why this framework seems promising is the mutual exclusiveness of the ontologies. Just as it is senseless to talk about a 'shark' (a type of 'entity') being 'two hours long' (predicates for a 'process'), so sub-ontologies are mutually exclusive (it is senseless to talk about a 'shark' that 'wilts' or a 'flower' that 'bites'). In the same way, 'deterministic' events are incommensurate with randomness. Other events, assigned into the 'random' ontology, are forbidden to have deterministic properties. We find that this framework can thus account for a stance of 'random is incompatible with predictable'. As a result of this mutual exclusiveness, the ontology framework argues that trying to help students learn a concept located in one ontology by drawing on their ideas in a different ontology would likely only lead to confusion.(Chi & Slotta, 1993; Slotta & Chi, 2006) For example, Slotta and Chi write

> Teachers should not try to "bridge the gap" between students' misconceptions and the target instructional material, as there is no tenable pathway between distinct ontological conceptions. For example, students who understand "force" as a property of an object cannot come gradually to shift this conception until it is thought of as a process of interaction between two objects. Indeed, students' learning may actually be hindered if they are required to relate scientifically normative instruction to their existing conceptualizations. (p. 286 of Slotta & Chi, 2006)

Methodology

This paper is a literature review of naïve ideas published in papers in physics topics related to probability. The topics investigated were decided via a Delphi study, which we will now describe. Following that, we will describe how we selected the naïve ideas within each topic to include in this literature review.

Delphi study

A Delphi study is a means of interviewing a group, often of experts, via several rounds of surveys. Unlike group discussions conducted in person, this method enables participants to remain anonymous. This ensures, for example, that more introverted experts have space to voice their opinions without the discussion being dominated by extroverted experts. The second author conducted a Delphi study with the goal of building consensus among the participating experts (Häder, 2009) via a website where respondents should rate different topics according to their relevance to probability. The authors primarily utilised the physics section of the International Baccalaureate Curriculum (IBO, 2014) compilation of topics and associated subtopics for the study. Several subtopics from the compilation were not included in the study, as they were deemed to be mathematical without a specified physical context (e.g. 'Vectors and Scalars'), historical in nature (e.g. 'The Beginnings of Relativity'), or specific applications of physics topics (e.g. 'Medical Imaging'). In addition, we added two topics, Non-linear Systems and Elementary Particle Physics, as we were interested in how respondents would perceive the relationship of these two topics to probability. This resulted in eleven topics, each consisting of two to six subtopics, as shown in Table 1. The website instructed the participants of this study to rate these subtopics according to their relevance to a literature review on probability-related physic topics. Figure 1 shows an example item.

Since our Delphi study intended to answer the question 'Which physics topics are relevant in terms of probability?', we needed experts with a broad knowledge of physics. We chose physicists at CERN as our target population. Although one might think at first that only particle physicists work at CERN, there are actually a large number of employees specialising in theoretical physics, accelerator physics, thermal physics, nuclear physics, and technical physics. Prior to the main study, the second author invited three physicists and three physics educators to post responses to the website as a pilot study. The participants needed about 10 minutes to finish. All respondents found the topic selection complete and well structured. Following the successful pilot study, the second author contacted twelve CERN physicists for the field study, of which ten completed and submitted the form.

In Delphi studies, a so-called 'definition of consensus' of 80% is common. (Diamond et al., 2014) In the case of this study, that meant that if eight or more of the experts vote for or against relevance of a subtopic, then the experts have reached consensus, and no follow-up voting on these subtopics is necessary. We determined consensus by collapsing the six-point scale into a two-point scale. If the experts responded to the statement that a subtopic should be included in the literature analysis with 'I strongly agree', 'I agree' or 'I agree slightly', we coded the response as 'vote to include'. Otherwise, the response was coded as 'vote to NOT include'. In the first round, experts reached consensus on 27 of the 43 subtopics, as shown in Table 1, and so we removed these subtopics from the second round.

Following the first round, the second author informed the experts about the 27 subtopics for which consensus had been obtained, and asked them a second time to judge whether or not the remaining 16 subtopics should be included in the literature review. She also encouraged the experts to provide comments. After this round, experts agreed not to include the subtopics 'Motion', 'Forces', and 'Momentum and Impulse', implying that the topic of 'Mechanics' should not be included in the literature review.

Table 1. The topics and subtopics of the Delphi study. Red shading indicates consensus (at least 8 out of 10 in agreement) after the first round that the subtopic should not be included. Green shading indicates consensus that it should be.

	I strongly agree	I agree	I slightly agree	I slightly disagree	I disagree	I strongly disagree
Mechanics						
Motion	1	1	2	3	2	1
Forces	1	1	1	3	3	1
Work, energy and power	1	1	0	4	3	1
Momentum and impulse	2	1	0	4	2	1
Circular motion	0	0	1	4	4	1
Newton's law of gravitation	0	1	0	4	4	1
Measurements and uncertainties						
Measurements in physics	6	4	0	0	0	0
Uncertainties and errors	8	2	0	0	0	0
Non-linear systems						
Non-linear dynamics	2	4	2	2	0	0
Chaos	4	4	1	1	0	0
Thermal physics						
Modelling a gas	3	5	2	0	0	0
Thermodynamics	6	3	1	0	0	0
Thermal energy transfer	1	6	1	1	1	0
Atomic, nuclear and particle physics						
The interaction of matter with radiation	5	4	1	0	0	0
Nuclear physics (including nuclear reactions)	4	5	0	1	0	0
Quantum physics	7	3	0	0	0	0
Discrete energy	2	4	2	2	0	0
Radioactivity	5	3	1	1	0	0
Wave phenomena						
Simple harmonic motion	1	1	3	2	2	1
Single-slit diffraction	4	1	3	0	1	1
Interference	3	2	3	0	1	1
Resolution	1	2	3	2	1	1
Doppler effect	0	0	1	6	1	2
Electricity and magnetism						
Electric fields	1	0	1	3	3	2
Magnetic fields	1	0	1	3	3	2
Magnetic effects of electric current	1	0	1	3	4	1
Electromagnetic induction	1	1	0	3	4	1
Relativity						
Lorentz transformations	0	2	0	3	3	2
Spacetime diagrams	1	1	0	3	3	2
Relativistic mechanics	0	1	2	2	3	2
General relativity	0	1	3	3	1	2
Astrophysics						
Stellar quantities	1	0	2	5	2	0
Stellar characteristics and stellar evolution	0	3	0	5	2	0
Cosmology	2	1	3	3	1	0
Stellar processes	0	3	0	5	2	0
Engineering physics						
Rigid bodies and rotational dynamics	0	0	2	3	4	1
Fluids and fluid dynamics	1	1	1	3	3	1
Forced vibrations and resonance	0	1	1	5	2	1
Elementary particle physics						
Fundamental interactions and forces	2	4	0	3	1	0
Structure of matter	0	3	4	2	1	0
Particle decay	5	4	1	0	0	0
Higgs mechanism and higgs boson	3	1	3	2	1	0
Standard model	1	4	2	2	1	0

Delphi Study on probability-related physics topics

Page 5 of 12

Thermal physics

The next topic we are looking at is thermal physics. Please rate the following statement:

"The following subtopic of thermal physics should be included in a literature review about probability-related topics in physics." *

	I strongly agree	I agree	I slightly agree	I slightly disagree	I disagree	I strongly disagree
Modelling a gas *	○	○	○	○	○	○
Thermodynamics *	○	○	○	○	○	○
Thermal energy transfer *	○	○	○	○	○	○

< Previous Page   Next Page >

Figure 1. Example item from the Delphi Study.

Correspondingly, we do not discuss naïve ideas in mechanics below. In the same way, experts agreed that the topics of 'Electricity and Magnetism', 'Relativity', and 'Engineering Physics' should not be included. On the other hand, experts agreed that the following topics should be included in a literature review concerning probability, and we discuss them below:

- Measurements and Uncertainties
- Thermal Physics
- Atomic, Nuclear and Particle Physics
- Non-linear systems
- Elementary Particle Physics

Finally, while experts agreed that the subtopics of 'Single Slit Diffraction' and 'Interference' within the topic of 'Wave Phenomena' should be included, they also agreed that the subtopics of 'Doppler Effect' and 'Simple Harmonic Motion' should not be included. From the comments of the experts, it became clear that the reason for inclusion of 'Single Slit Diffraction' and 'Interference' is that the experts saw a connection to quantum physics and therefore considered the topics as related to probability. We accounted for this by including in our literature review naïve ideas regarding quantum physics. Finally, out of the subtopics in Astrophysics, experts reached consensus on only one, 'Cosmology', that it should be included. Consequently, we have a section of our literature review examining naïve ideas in cosmology.

Collection of naïve ideas

When possible, we utilised two databases of student naïve ideas, Ideas Previas (Flores et al., 2014) and the database compiled by Duit. (Duit, 2009) The Ideas Previas website, under various headings like 'Structure of Matter' (which contains naïve ideas related to Quantum Physics; Atomic, Nuclear and Particle Physics; and Elementary Particle Physics), catalogues more than 1,000 naïve ideas documented in physics education research. Each naïve idea has its own short descriptive title like 'The half-life of a radioactive isotope is the period in which [it] loses its radioactivity. During this period [it] is dangerous', together with a citation for the source which documented the idea. In addition to this 'Structure of Matter' heading, we considered also the list of naïve ideas under the heading 'Thermodynamics', and 'Astronomy' (from which we found naïve ideas related to cosmology). We will describe below how we decided which naïve ideas from these lists to include in this literature review. As the Ideas Previas website does not contain naïve ideas relating to Non-linear Systems and Measurements and Uncertainties, we took a different approach for these topics. Namely, for Non-linear Systems, we used the main source for the Ideas Previas website, Duit's database of articles documenting naïve ideas. (Duit, 2009) The Duit database contains 23 journal articles related to student naïve ideas about Non-linear Systems. All of these articles were read by either the first author (if the articles were in English) or the second author (if the articles were in German), and naïve ideas from the articles were gleaned for inclusion in this literature review. Since even the Duit database does not contain articles related to student naïve ideas about Measurements and Uncertainties, we utilised yet a third method. Here, we sought out recent publications by researchers prominent for their studies of naïve ideas in the field of measurement in English (Saalih Allie) and in German (Susanne Heinecke). We then read the citations of these papers that 1) were either in English or in German and that 2) seemed relevant in that they potentially related to student difficulty with probability. We then read the citations of those cited papers, and so on, until no additional citations met these two criteria.

Selection of naïve ideas from ideas previas

All three authors coded all naïve ideas on the Ideas Previas website under the headings of 'Structure of Matter' and 'Thermodynamics' for their potential relevance to this literature review. The coding scheme was initially generated by examining a small subset of the naïve ideas taken from various physics topics from the Ideas Previas website. After discussing this small subset to calibrate coding, all three authors then individually coded all the naïve ideas under the heading of 'Thermodynamics' for whether we felt

that the idea relates to probability. With duplicates and correct ideas removed, this totalled 273 naïve ideas coded. In our coding, we decided to err on the side of having too many articles compared to too few. As such, we would read all articles that, after discussion, at least one of us thought might be relevant to our study, and so we coded for whether or not we thought the naïve idea should definitely *not* be included in the study. There were 174 disagreements, meaning that one or two authors had coded 'no' that the idea does not relate to probability, but not all three authors. This corresponds to 38.5% agreement (κ = 0.18). After discussion of these disagreements, we reached 100% agreement regarding which naïve ideas to definitely exclude from the literature review. Furthermore, we reached a considerably high consensus of 97.4% (κ = 0.97) regarding which naïve ideas definitely *should* be included. During this discussion, we decided to include any naïve ideas that consider both a microscopic and macroscopic view, as they would reflect probability-related confusion. In the previous case of the pressure increasing with the compression of an ideal gas, for example, we could envision a student with the naïve idea that each molecule would always exert an increased force on the container walls. Such an approach of attributing properties of the collective whole (here, pressure) to the individuals making up that whole would indicate a failure to understand the probabilistic description of the gas. This heuristic (that the naïve idea is to be included if it considers both a microscopic and macroscopic picture) led to the inclusion of many naïve ideas that did not explicitly discuss probability, such as 'At a very high altitude over the sea level air molecules are far less numerous, therefore pressure at such altitude is considerable less than pressure at the sea level' [sic]). We likewise decided that naïve ideas which do not consider both views (like 'The hot water is heavier because its volume expanded') are to be excluded.

After coding the thermodynamics naïve ideas, we repeated the process with the 127 naïve ideas (removing again duplicates and ideas marked as 'correct') listed under the 'Structure of Matter' heading on the Ideas Previas website. In this second round, all three authors again coded each idea independently for whether we felt that the idea should definitely be excluded from this literature review. In doing so, we bore in mind the two conditions discussed from the coding of thermodynamics-related naïve ideas (all naïve ideas considering both a microscopic and a macroscopic view are to be included; those which do not consider both views are to be excluded). Agreement (meaning that all three authors coded 'no, the idea should definitely not be included') was initially quite low, at 29.9% (κ = 0.07). After a prolonged discussion, we decided that naïve ideas pertaining to safety (like 'X-rays machine is based in that radioactive rays or waves go through a body without harm, and it projects an image of internal hard parts.' [sic]) do not necessarily relate to student reasoning about probability. As such, merely relating to safety was insufficient criteria to be included. We also decided that, in the minds of students, radiation is not necessarily at a smaller level than the object with which it interacts. Consequently, we decided to exclude statements like 'If a radioactive substance is near wood. It would get into the wood, some things are affected more than others.' [sic] As a result of these and other smaller points of discussion, agreement rose to 100% regarding which naïve ideas should definitely be excluded. We also reached a considerable consensus of 96.8% (κ = 0.96) regarding which naïve ideas should definitely be *included*.

In summary, the three of us reached agreement that, out of the 273 naïve ideas in thermodynamics, 186 do not satisfy the criteria described above, and hence are not to be

included. Twenty-one of the remaining 87 naïve ideas were cited as coming from a total of 5 papers written in Spanish, and this language barrier prevented us from reading them. We then read the journal articles from which the remaining 66 naïve ideas were taken. We discovered, for 26 of these 66 ideas, discrepancy between the categorisation of Ideas Previas and the articles corresponding to the ideas. Namely, either the article did not discuss the idea reported by Ideas Previas at all, or the article used the idea as an example of an answer considered to be correct. For example, one of the 66 ideas from Ideas Previas that we decided meets our criteria was 'Heat makes the molecules that are in [the material] … to move faster … causing more pressure because they bump against each other.' (Hewson and Hamlyn, 1984) The article, however, uses this student statement to demonstrate part of the (canonically correct) kinetic gas theory. Since it is not used in the article as an example of a naïve idea, we of course do not include it here in this literature review either. The remaining 40 naïve ideas that are included in this literature review are in Table A1 in the Appendix. Out of the 124 naïve ideas in Structure of Matter, the three of us agreed that 67 are not relevant, and 15 of the remaining 57 are from articles written in Spanish. We read the journal articles from which the remaining 42 naïve ideas were taken. Here again there were inconsistencies between Ideas Previas's categorisation and the descriptions within the corresponding articles, and this led to the removal of 14 ideas. The remaining 28 naïve ideas that are included in this literature review are in Table A2 in the Appendix.

Ideas Previas catalogues naïve ideas related to Cosmology under the heading of 'Astronomy'. In total, there are 118 naïve ideas under this heading. After individual coding and discussion similar to that just described, we decided that 102 of them are not related to Cosmology, as they do not explicitly relate to the formation of heavenly bodies (e.g. 'The night occurs because the Sun goes behind hill.' [sic])[4] Of the 16 cosmology-related ideas, we decided that 8 of them would not be relevant for this paper, as they do not plausibly relate to probability in some way. The remaining 8 naïve ideas are presented in Table A3 in the Appendix.

To recap, with the exception of the topic Measurements and Uncertainties, the papers we analysed came exclusively from Ideas Previas and the Duit database. As these catalogues only span until 2007, our literature review for these topics does not include more recent sources. As such, we cannot claim that our review is exhaustive. That is not, however, necessary for our purpose of illustrating that many naïve ideas in diverse fields of physics are due, to varying extent, to student difficulty with probability. As we will elaborate upon in the Discussion section below, we are in fact aware of additional naïve ideas that our methodology did not uncover, but that nevertheless could be included in a more exhaustive review of probability-related naïve ideas.

While not necessary for our argument, we did nevertheless hope that our literature review would be representative of the range of probability-related naïve ideas in physics. We felt justified in the decision to use Ideas Previas and the Duit database with the hypothesis that the topics within had been extensively researched by the point at which the databases stopped collecting data, and that findings had approached saturation. We checked this hypothesis by the following means: first, we searched Google Scholar with the Boolean string 'allintitle: thermodynamics student OR students OR conception OR misconception' and the requirement that the year be 2008 or later. This yielded 151 articles. The first and second authors read the titles of the articles and the corresponding abstracts. In

total, there were 53 articles that at least one of us felt might potentially contain naïve ideas that are related to probability. Thirteen of these articles were not readily accessible, but the first author randomly chose and read ten articles from the remaining 40. Of these ten, three contained naïve ideas that plausibly were enabled by student difficulty with probability: Sriyansyah et al. found that, given two processes, one in which much work is done and one in which little work is done to reach the same ending state, students think that the internal energy is different, depending on the process. (Sriyansyah et al., 2015) Saricayir et al. discussed how students struggle with recognising the role that mass plays in the heating and cooling of objects. Students conflate heat capacity and specific heat, or think that a glass of hot water has more internal energy than the ocean. The researchers also noted student difficulty with heat transfer between two objects, thinking that equilibrium is when the internal energies of the two bodies are equal (as opposed to the two temperatures). (Saricayir et al., 2016) We see the naïve ideas from these two papers as being consistent with difficulty in distinguishing heat from temperature, which we found in the Ideas Previas database and which we discuss below. Finally, Bowen et al. documented how some students try to use the ideal gas law or the first law of thermodynamics to reason about a single molecule getting hit by a moving piston in a sealed box. (Bowen et al., 2015) This naïve idea is rather distinct from what we discuss in this literature review. In summary, from our random sampling of recent papers in Thermal Physics, we found only one naïve idea not captured by what we review below, consistent with our hypothesis that our review is representative of probability-related naïve ideas.

Results

Here we present the results of our literature search. We begin with a review of naïve ideas relating to Non-linear Systems from articles listed in the Duit database. (Duit, 2009) We then turn our attention to naïve ideas in Thermal Physics; Atomic, Nuclear and Particle Physics; Elementary Particle Physics; Quantum Physics; and Cosmology from articles listed in Ideas Previas. (Flores et al., 2014) We end with a review of articles pertaining to naïve ideas in Measurements and Uncertainties.

Non-linear systems

Although much formal science education involves linear systems (where doubling x results in doubling y), most everyday experiences are non-linear in nature. Many non-linear systems, despite being deterministic, are nevertheless unpredictable because infinitesimally small perturbations to the initial state give rise to dramatically different outcomes. (Boeing, 2016) The classic example is to consider the flapping of a butterfly's wings giving rise to a hurricane on the other side of the world. Meteorologists do not expect to account for the flapping of every butterfly's wings. Nevertheless, they are able to forecast the weather by running simulations that combine a range of potential perturbations, treated as being random effects (in the sense that they cannot be predicted), with deterministic laws of how air will move when faced with temperature and pressure gradients. Randomness must be combined with things which are predictable in diverse contexts ranging from the migration of birds, to the build-up of traffic jams, to the response of the public to a new government regulation.

Although relatively little research has been conducted on student understanding of non-linear systems, the studies that have been published suggest that people have a tendency to either think that things happen by chance or that they are predictable, and not that deterministic laws could lead to unpredictable outcomes, depending sensitively upon initial conditions (*e.g.* Stavrou & Duit, 2014). Some students, for example, display so-called 'Laplace determinism', the view that all processes are deterministic. (Buecker et al., 1999; Komorek, 1998; Stavrou et al., 2003, 2005; Stehlik & Pade, 2006) Several researchers observed student difficulty with various typical chaotic systems, such as a chaotic pendulum, and found the student view that if a situation can be predicted with outcomes that vary only slightly, then the situation is deterministic. If not, it is random. (Komorek, 1998; Stavrou et al., 2008) Some students assume that if the starting conditions are similar, then the results must be similar as well or, conversely, that similar effects must have similar causes. (Buecker et al., 1999) In looking at dendrite formation, students either expect the same number of branches to form the second time because they think there is no randomness in science, or they think that the structure is entirely random, and they are then surprised to see that it is always dendritic. (Komorek et al., 1999, 2001) Similarly, in situations where both static and dynamic descriptions of a situation coexist, such as in modelling population stability, students have a tendency to think exclusively in a static way. (Bell & Euler, 2001) These naïve ideas can be described well with the ontologies framework. Namely, we can account for them with the argument that students assign some events (where outcomes of repeated trials vary only slightly, for example) into the 'deterministic' ontology. As such, these events could not have characteristics of randomness. 'Laplace determinism', for example, could be described as the learner assigning *every process* into this ontology. Other events (where the outcomes vary more dramatically), assigned into the 'random' ontology, would be forbidden to have predictable properties.

Non-linear systems are often visualised with fractal images. One type of fractal image that can be found approximately in nature, for example, in fern leaves, is formed by repeating a system rule over and over again (*e.g.* the Serpiensky triangle follows the rule of starting with a triangle and recursively subdividing into smaller equilateral triangles). As one might expect, students find the process of fractal formation un-intuitive at first. As such, when shown the first three steps in the formation of the Koch curve[5] and asked to draw the fourth step, some students draw a star emerging from the line. Likewise, when shown a cylindrical electrode and asked to predict what shape of precipitate will grow on the electrode, students predict a cylindrical structure instead of the dendritic shape. (Komorek et al., 1999) When given images of ferns, etc. and asked to organise them into groups that make sense to them, students who have not yet learned about fractals do not focus on the aspect of self-similarity. At best, they pick up on the characteristic of being 'branched' or of being 'branched repeatedly'. (Komorek et al., 2001) Thinking in terms of the 'deterministic' and 'random' ontologies, perhaps the circular electrode, being of a geometry frequently encountered in science and maths classes that prioritise problems posed as having a definite solution, promotes assignment of phenomena into the 'deterministic' ontology. From within this viewpoint, it is natural that one would expect that a pattern should emerge which exhibits circular symmetry.

Some research concerns instruction aiming to equip students with general thinking strategies for reasoning about non-linear systems in general, independent of the specific situation. However, students tend to focus on superficial features instead of building

awareness of the key ideas, such as that a minor perturbation can give rise to vastly different outcomes. (Bell & Euler, 2001; Wilbers & Duit, 2001) When asked to describe the motion of a chaotic pendulum with two magnets, for example, students explained the deviation of the pendulum from the plane of symmetry as arising from the rotation of the Earth, likening it to Foucalt's pendulum, seemingly ignoring the fact that sometimes the pendulum ends up near one magnet and sometimes the other. (Duit et al., 1998) Some researchers have argued that part of the difficulty is a linguistic one. Apparently, for some students, the phrase "chaotic system" is an oxymoron, as "system" denotes a sense of order, which therefore cannot be chaotic. (Roth & Duit, 2003) Students generally associate the word 'order' with things that are static, like a clean room or a balanced see-saw (Bell, 2007) or with things that are predictable. (Stehlik & Pade, 2006) Although students are familiar with the word 'feed-back' from music, they do not associate it with dynamic balance found prevalently in chaotic systems. (Bell, 2007) It is also plausible that prior education in maths and physics classes, where topics that are deterministic and well-ordered are the focus, make it difficult for students to grasp chaotic systems. In the case of fractals, Vacc wrote that children who had previously studied geometry had a harder time describing the shapes of fractal-shaped objects like popcorn and kale leaves. (Vacc, 1999) In terms of the ontologies framework, these naïve ideas can also be explained with the argument that students assign some things (like the pendulum in the anecdote above) into the 'deterministic' ontology and other things into the 'random' ontology. This dichotomous categorisation scheme is encouraged by language use and typical classroom experiences and it makes the core ideas of non-linear systems untenable.

Cosmology

We will next review naïve ideas discussed in literature specific to the field of Cosmology, continuing to utilise the ontologies perspective. Like with Non-linear Systems, random-ness is intertwined with predictable outcomes in Cosmology. The normative view utilises the nebular theory (that is, that most heavenly bodies form from the coalescing of gas and dust). At a microscopic level, it is considered random which areas of space will contain what concentrations of cosmic dust. However, it is nevertheless predictable that, given a critical mass of this dust in a given location, that coalescence will occur. Education researchers have documented several student naïve ideas in this field, and we again find the ontologies framework productive for thinking about their commonality.

Marques and Thompson found that many students think that nothing existed before the Earth and that, for example, it was formed immediately after the Big Bang. Other students attribute spiritual significance to the beginning of the Earth, thinking, for example, that there was 'something like the Garden of Eden' before the Earth was formed. Other students in their study, reflecting exposure either formally or informally to scientific ideas on the subject, described the presence of 'vacuum' with a few small particles and 'fire' before the Earth was formed. (Marques & Thompson, 1997) Some students describe the Earth originating by the collision of planets (Marques & Thompson, 1997) or rocks. (Sharp, 1996) Students attribute collisions, for example, of the Sun with 'something', to the formation of other planets as well. (Sharp, 1996) Sharp found that some students, in discussing the Big Bang, described everything in the universe 'whooshing out' when the gases inside exploded. (Sharp, 1996) From the perspective of the ontologies framework,

we can account for at least some of these naïve ideas by noting how they deviate from the nebular theory. A student assigning the formation of a stellar body into the 'deterministic' ontology would be expected to reject the idea of dust accumulation, as it inherently involves randomness. Instead, we would expect such a student to attribute intentional (for example, spiritual) design to the Earth's creation or to assume that it was always present. Given exposure to the idea of the Big Bang, we might think such a student would imagine the Earth to be pre-formed inside the 'big lump' to come rushing out when the gases exploded.

Atomic, nuclear, and particle physics; Elementary particle physics; And quantum physics

Although the time at which an individual radioactive nucleus fissions is taken to be random, when enough of these atoms are collected, one can predict how long it will take for half of them to fission (the 'half-life'). Such a phenomenon would not be conceptually accessible for a student who is assigning the process into the 'deterministic' ontology. We should expect, instead, that such a student would think that it is also predictable when individual nuclei will fission – for example, that *all* atoms will have fissioned after one half-life. This is consistent with research findings on student naïve ideas, for example, that the material will be safe and no longer radioactive after one half-life. (Lijnse et al., 1990) It could be the case that students similarly conflate radioactive material and radiation (Alsop, 2001; Millar & Gill, 1996) as a result of assigning the effects of radiation on the human body into the 'deterministic' ontology. The normative view is to treat it as random whether or not a piece of ionising radiation will interact with DNA, and it is further considered random whether or not that ionisation event will lead to cancer. When many of these individual events take place, however, we can predict the outcome: sufficient over-exposure to ionising radiation leads to death. Such a two-level picture where randomness prevails in the microscopic level is likely difficult for students to come to terms with. It is easier to label the ionising radiation as 'bad' or 'dangerous' and, by extension, the radioactive material likewise. Using the 'deterministic' ontology, learners may group radiation and radioactive material together as a 'poison' in this way. An additional result of this conflation could be decreased importance given in the minds of learners to the origin of this 'poison'. Rather than recognising nuclear energy as arising from interactions between constituents of the nucleus, students instead describe it as being released as a result of collisions between nuclei, atoms, molecules, or 'particles'. (Cros et al., 1988)

As just discussed, students attribute characteristics of the whole to the constituents comprising that whole when reasoning about radioactivity (for example, that an individual atom has a half-life, implying that all the material will lose its radioactivity at the same time. (Lijnse et al., 1990) Education researchers have found the same tendency to attribute the macroscopic characteristics to the microscopic level in other areas of atomic physics as well. For example, when asked to describe the internal structure of metal, many students described it as looking like lamina or little stones, things that one might expect to see if one would cut the metal in half and look with the naked eye. Although such a description could arise from students not understanding what was meant by 'internal structure', other students described the atoms of a metal as being packed more closely

together than they are for glass because some space must exist in the glass for light to be able to pass through. (De Posada, 1997) Other students have described the atoms of solids as being harder than the atoms of liquids. (Harrison & Treagust, 1996)

In using the analogy of an electron 'shell', we attempt to explain to students that, although it is random where the electron in an atom will be found should a position measurement be made, the energy will predictably be that corresponding to its shell. A student who either 1) focuses on the predictable results of energy measurements and assigns the atomic configuration into the 'deterministic' ontology, or 2) focuses on the random location of the electron and hence uses the 'random' ontology is likely to have difficulty comprehending this analogy and using it effectively. Education research has indeed found naïve ideas regarding this and other analogies. For some students, the metaphor conjures an image of protection. They hence think that electron shells protect the nucleus in much the same way that egg shells protect the egg. (Harrison & Treagust, 1996) Other students think of the shell as having a fixed radius that confines the electrons, and that it is this rigidity that prevents electrons from falling into the nucleus under the electrostatic attraction. (Fischler & Lichtfeldt, 1992) Similarly, when learning about an electron 'cloud', students think of a material within which raindrops are embedded and attribute a material-like property to the electron cloud itself, as a sort of stuff within which electrons are embedded. Accordingly, students attribute electrons in this cloud with the property of being very small (like a raindrop). In fact, although electrons exhibit point-like behaviour when outside of an atom (for example, when boiled off an electrode in an electron gun), within an atomic electron cloud, it is more appropriate to model the electron as being a charge distribution. A dramatic example of metaphor and language-related difficulties is students thinking that the nucleus of an atom contains instructions for the atom to follow and is responsible for the division of the atom. Apparently, these students are confusing the nucleus of an atom with the nucleus of a cell. (Harrison & Treagust, 1996) Many students think that atoms look like the figures used to represent them. (Griffiths & Preston, 1992; Hogan & Fisherkeller, 1996) When asked to 'sketch what you would see if you could take an atom and look at it under a microscope so powerful that all the details of the atom could be seen', for example, many students drew a sphere with components inside. (Griffiths & Preston, 1992)

Taber found that students struggle in understanding the electrostatic forces involved in keeping electrons inside of atoms. For example, some students claimed that each electron in an atom, regardless of its distance from the nucleus, would experience the same force pulling it towards the centre. Some of these students further explained that the nucleus has a fixed amount of force with which it can attract all of the electrons, and that the electrons need to share this force equally. If one electron is removed in an ionisation event, it follows that the remaining electrons will each have a larger share of the force on them. This faulty model is used even by students who demonstrate mastery of the use of Coulomb's law on traditional physics questions with stationary charges. (Taber, 1998) We find it plausible that one reason why the student model is popular in this context is that students are assigning the situation into the 'deterministic' ontology. Namely, faced with predictable outcomes of various changes (ionisation energy increases with successive ionisation events, atomic radius generally increases as the number of electrons does), students create a simple and predictable scheme leading to those results. At the same time, we speculate that Coulomb's law is relatively inaccessible to students in this context, because it relies upon positions of

charges, which have inherent randomness here. For a student assigning the physical location of the electrons into the 'random' ontology, Coulomb's law, which is conceptualised as a tool to make predictions, would not be applicable.

Thermal physics

Above, we discussed the naïve idea that transparent solids have atoms that are more spread out so as to allow the light to pass through. (De Posada, 1997) An analogous naïve idea in the field of Thermal Physics has been documented, that materials in which heat travels faster have atoms that are less densely packed. (Hewson & Hamlyn, 1984) We find it plausible that both naïve ideas are fostered by assigning the process (light/heat passing through) into the 'deterministic' ontology. We can imagine a student argument such as 'we know that heat will predictably travel better from one end to the other if it is metal than if it is wood, so there must be some difference inside that allows for that to happen. Maybe it is like water flowing in a pipe – the less clogged the pipe is, the better the flow. The only thing inside the material is atoms, so less clogged means that the atoms are less tightly packed together. The atoms predictably are located at a greater distance from each other.' As with this example, many naïve ideas in Thermal Physics involve students erroneously attributing an aspect of predictability found in the collective to the individual elements that comprise that collective.

Temperature and heat transfer

Life is full of everyday experiences of warming and cooling. If we imagine a learner assigning these processes into the 'deterministic' ontology, we would imagine an account like the following: 'Some objects are hot and others are cold, and they interact with each other. If a hot object has some structure (is made up of parts), then those parts are hot as well, and likewise for cold objects.' For this user of the 'deterministic' ontology, the idea that some parts of the hot object are randomly less 'hot' than other parts does not seem tenable, as processes assigned into the 'deterministic' ontology do not have attributes of randomness. Students with this view are unlikely to see a need to distinguish between heat and temperature (Erickson, 1979; Hewson & Hamlyn, 1984; Kesidou & Duit, 1993; Rozier & Viennot, 1991; Summers, 1983), as they think of heat (or cold for that matter (Erickson, 1979)) as being a substance (fumes or bubbles (Erickson, 1979) or warm wavy-air from a fire (Hewson & Hamlyn, 1984), for example) contained within and exchanged between objects. Within this (caloric) view of objects containing heat, researchers have found students explaining things like metal balls swelling when heated in terms of the ball holding more heat (Hewson & Hamlyn, 1984) and cooling as taking place when fumes previously contained within the object are released. (Erickson, 1979)

Learning a particle description seems to be insufficient for learners to relax this ontological assignment. Some students, for example, say that temperature is the speed of particles, which suggests that the particles have a uniform speed as opposed to a random distribution. (Kesidou & Duit, 1993) Other students continue to conflate heat and temperature, thinking, for example, that heat is the movement of particles (Kesidou & Duit, 1993) or the energy of moving particles. (Summers, 1983) We would expect students using the 'deterministic' ontology to argue that, if hot objects give heat to cold objects and both are made up of particles, then it must be the case that hot particles give heat to

cold particles. Indeed, this naïve idea has been documented by researchers. (Kesidou & Duit, 1993)

Although it is admittedly a more tenuous claim, it may be the case that assigning heat transfer into the 'deterministic' ontology contributes to the naïve idea that, if heat makes metal increase its size, then it follows that the individual particles are increasing as well, either in size or number. (Hewson & Hamlyn, 1984; Kesidou & Duit, 1993) Similarly, if heating is thought to predictably destroy an object (like burning wood, for example), then it may follow for some students assigning the process into the 'deterministic' ontology that, in the case of heating metal, the inner composition is being destroyed. (Hewson & Hamlyn, 1984)

Phase transitions

Although heat transfer is, in fact, predictable (we can assume an overall heat transfer from hotter objects to colder objects), it is driven by stochastic processes (in the case of conduction, for example, molecules of the hot object collide with molecules of the cold object, usually losing but sometimes gaining kinetic energy). Phase transitions in thermodynamics similarly are predictable on the macroscopic scale but stochastic on the microscopic scale, and we can similarly consider that students who assign these processes into either the 'deterministic' or 'random' ontologies might exhibit naïve ideas here as well. Research has indeed documented naïve ideas pertaining to phases and their transitions. While some students may succeed in reciting that condensation on the inside of a sealed bottle containing a bit of water is driven by a temperature difference without being able to justify *how* that happens (Chang, 1999), others may mis-memorise the statement as being due to a pressure difference or argue that condensation arises when something encounters heat. (Chang, 1999) It has been found that some students think condensation is a process of air (instead of water vapour) turning to liquid water (Chang, 1999; Ure & Colinvaux, 1989) and vice versa for evaporation. (Ure & Colinvaux, 1989) Regarding condensation on the outside of a glass of cold water, some students think that this water comes out through the glass walls and still others think that the water comes from within the glass itself and is extruded as sweat. (Ure & Colinvaux, 1989)

Research has found other naïve ideas from learners that can perhaps be considered to be more advanced in their learning. Many learners explain that the phase of an object depends only on temperature because the particles decrease in motion (which cannot explain why a solid glass and liquid water within it can be at the same temperature). (Rozier & Viennot, 1991) Some students attribute an important role to the air above an open bottle of water as facilitating the evaporation of the water molecules by allowing them to combine with the air molecules (Chang, 1999) and some think that solute particles evaporate together with the water. (Hwang & Hwang, 1990) A student assigning phase transitions into the 'deterministic' ontology, for example, might be imagined to reason 'if there is less salt water than there was before, then the salt water must have turned into a gas'. Lacking from the argument is the microscopic picture of water molecules that are bonded together less tightly than salt molecules. It is plausible that one reason why such a microscopic picture is lacking is that the microscopic picture is random in nature, and hence does not belong with a predictable phenomenon like water evaporating.

Ideal gases

In the previous section, we conjectured that the stance 'random is incompatible with predictable' might be part of the cause for students failing to map the microscopic stochastic picture of particles colliding and bonding with the macroscopic predictable picture of changes of state. If this conjecture is correct, then we should expect to find additional naïve ideas in the context of ideal gases, where several such mappings must take place simultaneously. Rozier and Viennot documented naïve ideas in this topic. In an adiabatic (that is, no heat exchanged) compression of an ideal gas, students explained that the pressure increases because the volume is decreasing, resulting in a greater collision density. Although it is indeed true that the collision density is increasing, what these students fail to discuss is that the temperature of the gas is also increasing during the compression (this is due to the work done on the box. Gas molecules that collide with the moving lid will rebound with a greater kinetic energy than if the lid were stationary, in which case the collision is elastic). When asked specifically about the temperature of the gas increasing, rather than the correct explanation (of work being done by the lid), students use the increasing pressure as an intermediate step in their explanation. For example, they argue that there is 'an increase of temperature because there are more collisions between particles [greater pressure] and more energy is produced due to friction [increasing the temperature].' (Rozier & Viennot, 1991) This naïve idea was also observed by Manthei. (Manthei, 1980) The fallacy of this argument, as Rozier and Viennot point out, is that the ideal gas model assumes elastic collisions between molecules. Indeed, the authors argue, if such collisions produced heat, then a gas kept in an adiabatic container would soon explode – steady-state conditions would be impossible. Related to this is the naïve idea that the temperature of a gas in a container being heated increases less when the box is bigger and the (same number of) particles are hence more diffuse (Rozier & Viennot, 1991) – students with this idea seem to be conflating pressure (and the corresponding collisions between particles) with temperature. An additional example of failing to bear in mind multiple macroscopic quantities at a time while reasoning with a particulate picture is the idea that the pressure decreases as altitude increases as a result of there being fewer air molecules and hence fewer collisions (the failure is again in noting that the temperature also drops with altitude). (Rozier & Viennot, 1991)

Summary: student use of ontologies in reasoning about Thermal Physics

Situations where a stochastic description must be combined with a predictable description abound in Thermal Physics. One situation discussed was heat conduction, regarding which researchers have found the naïve idea that hot particles give heat to cold particles. (Kesidou & Duit, 1993) This naïve idea is a particularly clear example of our argument, that students struggle in part because of assigning a given process into either the 'deterministic' or 'random' ontology. In this case, if learners assign heat conduction into the 'deterministic' ontology, then we should not be surprised that the idea that it is random whether an individual atom gains or loses energy in a collision does not make sense. It is plausible, we argue, that assignment into these two ontologies causes difficulties for students in less obvious situations where a macroscopic description must be connected with a microscopic description as well. For example, we can envision naïve ideas such as 'condensation of water on the surface of glass occurs as a result of a pressure difference' (Chang, 1999) to be the result of a student who tried to memorise by rote and ended up

memorising incorrect information. Although a student might attempt to memorise without understanding for a number of reasons, we would expect it to be more common in situations where the explanation (in this case, 'the motion of water molecules is random, but they have a distribution that changes in a predictable way when the water is heated or cooled') is untenable to the student. This would surely be the case for a student who thinks that things are either predictable OR random, but NOT BOTH at the same time.

To be clear, we do not claim that categorising with ontologies is the ONLY cause of the naïve ideas discussed above. With the naïve idea that heat is better conducted in metal than wood because the atoms are more spread out (Hewson & Hamlyn, 1984), for example, a correct explanation would require understanding that heat is conducted in metal via free electrons scattering off impurities and lattice defects, whereas in wood it is through collisions between adjoining molecules. Our argument is that, if we envision a student thinking of heat conduction from within the 'deterministic' ontology, we find it unsurprising that the student would propose the model that atoms inside of a metal must be more spread out to allow the heat to flow unimpeded.

Measurements and uncertainties

Like science in general, a key aspect of physics is its empirical nature, and hence measurement plays a paramount role. Physics teachers often assume that students will automatically come to understand the measurement and data analysis process in carrying out labs, but education research has shown that this is largely not the case. Even after having a number of physical science laboratory experiences, myriad naïve ideas remain. As with the previous sections, we will consider the utility of the ontologies framework in understanding these naïve ideas, specifically in service to the hypothesis that naïve ideas are enabled by students assigning phenomena into either the 'deterministic' ontology or the 'random' ontology.

Multiple measurements and taking an arithmetic mean

Education researchers looking at student difficulties with measurement have commonly documented students who do not see the need to take multiple measurements, unless, of course, the instructor asks them to do so, or something obviously went wrong with the first measurement. (Abbott, 2003; Allie et al., 1998; Buffler et al., 2001; Davidowitz et al., 2001; Gott & Duggan, 1995; Hackling & Garnett, 1992; Heinicke, 2012; Lubben & Millar, 1996; Millar et al., 1996; Millar & Lubben, 1996; Rollnick et al., 2001; Sere et al., 1993) As an example, when Cauzinille-Marmeche et al. had children test to see if container shape affects the time for a covered candle to extinguish itself, the children saw with the trial of two different containers that there was a difference in time. They did not repeat the experiment, however, as they did not expect that there could be a difference in time if the *same* container were used twice. (Cauzinille-Marmeche et al., 1985) A similar phenomenon is observed when students are presented with data. Given a single trial of a robot arm throwing ball A and ball B with the result that ball A travelled farther, third graders were confident that the next time the balls were thrown that ball A would again travel farther. (Masnick & Morris, 2002) Students such as these do not see a need to repeat their measurement because they expect the results to be identical, and are perplexed when the results are not. (Cauzinille-Marmeche et al., 1985; Coelho & Sere, 1998) Generally, research

has documented students collecting observations or making measurements with no apparent awareness of the uncertainty associated with the measurement process. (Deardorff, 2001; Ibrahim, 2006; Lubben & Millar, 1996; Warwick et al., 1999) This is exactly what one would expect if students are assigning the process of measurement into the 'deterministic' ontology. Namely, such a student might argue that measuring is a thing that predictably tells you information about the object being measured. Since it is predictable, it is not possible for there to be randomness involved, and hence every measurement would produce the same result.

A slightly preferable version of assigning measurement into the 'deterministic' ontology is the view that a measurement *should* (assuming all goes as expected) predictably yield the same (true) property of the object being measured. We might expect students with this view to recognise that readings will vary from measurement to measurement, and to attribute 'error' to this variance. However, when asked to identify such potential sources of error, we would not be surprised if these students tend to exclusively think about mistakes that they themselves might have made (a naïve idea documented by Goedhart and Verdonk (1991)), as they see the variation in the data as being due to personal mistakes (a view documented by Heinicke (2012), Leach et al. (1998), Pillay et al. (2008), and Rollnick et al. (2001)). We might further expect that such students would consequently repeat the experiment to get closer to a 'true value', thinking that they are increasing accuracy (whereas they are actually only increasing precision) (a naïve idea documented by Allie et al. (1998), Evangelinos et al. (2002), and Sere et al. (1993)). We might expect these students to thus report only one result despite taking many trials or to use only one data point to make a conclusion, an expectation confirmed by research. (Austin et al., 1991; Deardorff, 2001; Gott & Duggan, 1995; Lubben & Millar, 1994; Millar et al., 1994) Further aligned with what we would expect from students assigning measurement into the 'deterministic' ontology in this way, research has documented students who report only the last result, since that came after the most amount of 'practice' or who report the result that is a repeated value, since that is most likely to be the 'true value'. (Allie et al., 1998; Buffler et al., 2001; Davidowitz et al., 2001; Ibrahim, 2006; Lubben & Millar, 1994, 1996; Millar & Lubben, 1996; Pillay et al., 2008; Rollnick et al., 2001; Sere et al., 2001, 1993; Warwick et al., 1999) Research has shown that, even when presented a data set and asked 'what value to report as the final result', students answer with a single value (as opposed to reporting a range). (Buffler et al., 2001; Volkwyn et al., 2008)

Consistent with the student inclination to consider 'error' as comprised exclusively of experimenter mistakes, research has documented the student view that, given precise enough tools, a measurement with zero uncertainty can be obtained. (Deardorff, 2001; Heinicke, 2012; Ibrahim, 2006; Pillay et al., 2008) One student, for example, wrote 'The distance is exactly 423.7 mm, since electronic things always make no mistakes.' (Pillay et al., 2008) Perhaps because of the perception that less human judgement is required, Coelho and Sere found students with the view that if the distance measured lines up with the marking on the ruler, then it is a more precise measurement. On the other hand, they described other students who thought that the true value can be obtained in a measurement that is sufficiently COARSE, again, perhaps, because less human judgement is involved. (Coelho & Sere, 1998)

Still other students have been found who, despite taking multiple measurements, choose the first measurement (Lubben & Millar, 1996; Sere et al., 1993), for example, 'because the equipment was clean'. (Lubben & Millar, 1996) This raises the question of

why the student bothered to take multiple trials at all, with the likely answer being that students are taking multiple trials because they know that more measurements lead to a better result, but they are unsure about what exactly is better. (Lubben & Millar, 1994; Sere et al., 1993) One student, for example, explained the reason for repeating measurements to be 'you are meant to do an experiment three times' without feeling a need for further elaboration. (Lubben & Millar, 1996)

One reason to take multiple data points is to calculate an arithmetic mean. This is advantageous, as the arithmetic mean is generally more representative of the object being measured than an individual data point. If the object were to be measured an additional time, that measurement is most likely to be close to the mean than to, for example, a measurement that happens to have repeated. Researchers have found students who understand that it is important to calculate an arithmetic mean, but who do not understand why. For example, some students say that an experiment should be repeated with different parameters before an average is taken. Lubben and Millar write that this 'seems to arise from the student's attempt to reconcile the taught routine of repeating measurements to obtain a spread of values which can then be averaged, with the belief that repeating exactly the same measurement will yield exactly the same result.' (Lubben & Millar, 1996) Using the ontologies framework, it seems likely that these students, despite being aware of the importance of calculating arithmetic mean, have failed to make sense of the tool because of their assignment of measurement into the 'deterministic' ontology. If taking a measurement predictably yields the desired information about the object being measured, then it makes no sense to utilise arithmetic mean, as it is a tool designed to minimise the effects of randomness. Similar symptoms of students trying to memorise protocol without understanding is the misuse of the word 'average' (for example, to mean choosing a value that is repeated twice) or describing the correct procedure to find the arithmetic mean, but then abandoning that approach when presented with actual data that includes recurring values. (Lubben et al., 2001; Lubben & Millar, 1994) Finally, there are students who calculate the mean but then choose the data point closest to that mean to report their value, as demonstrated in the following quote. "They should not take the average because this does not refer to a result obtained from one of the experiments." (Buffler et al., 2001) Students seem to have a strong attachment to the idea that the true value is one of the measurements. (Ibrahim, 2006; Leach et al., 1998)

For some students, the importance of taking multiple measurements and taking an arithmetic mean is context-dependent. For example, although a student might not think to repeat a distance measurement, they might do so for time measurements (Allie et al., 1998; Heinicke, 2012) with an argument like 'it is tricky to measure time accurately with a stopwatch, so I reckon that you should take more than 2 readings. More readings would eliminate human error in stopping and starting the stopwatch when the average is taken.' (Allie et al., 1998) Although these students do seem to have an understanding of the benefit of the arithmetic mean, their decision to only repeat measurements where 'human error' is salient reflects the view, again, that a single measurement *should* (assuming all goes well) give you the true value of the object being measured.

Using data spread

Consistent with the naïve ideas just discussed, research has found students who ignore data spread after taking multiple trials and, when asked to provide a confidence interval,

report only the precision of the measuring instrument. Some students will hence report this 'confidence interval' for each trial. (Sere et al., 1993) When given two sets of data with identical means but different spreads and asked which set of data is better, many students answer that they are equally good because the averages are the same (*e.g.* Lubben & Millar, 1994) or because 'results are never wrong'. (Rollnick et al., 2001) Particularly this latter view is consistent with assignment of measurement into the 'deterministic' ontology. In addition, some students report that a larger spread represents a better measurement set (Lubben & Millar, 1994; Sere et al., 2001), which is consistent with the idea discussed above that different parameters should be used before taking the mean. The naïve idea that both sets are equally *un*trustworthy because 'every time you do it you will get a different answer' (Lubben & Millar, 1996) has also been documented. This can perhaps reflect students assigning measurement into the 'random' ontology. Namely, they might be thinking 'when I take a measurement it is random what value I get. Things that are random are not predictable. Hence, neither set of data helps me in predicting what the next measurement will yield/what the true property of the object being measured is.'

Just as students seem capable of taking multiple measurements and calculating an arithmetic mean without understanding why, so too can they make decisions involving data spread with a lack of conceptual understanding. When presented with the two data sets just described (same means, different spreads), many students report correctly that the set with the smaller spread is better. In spite of that, many of these students do NOT answer that way when the mean of the two sets is *different*. Instead, they now ignore the scatter of data and express dependence on a 'true value' or the uncertainty of individual measurements. (Heinicke, 2012; Lubben & Millar, 1996; Sere et al., 2001) Similarly, when asked whether two data sets with different means but overlapping spreads agree with each other or not, many of these students ignore data scatter and just use the arithmetic means (Abbott, 2003; Allie et al., 1998; Cauzinille-Marmeche et al., 1985; Davidowitz et al., 2001; Deardorff, 2001; Ibrahim, 2006; Leach et al., 1998; Lubben et al., 2001; Lubben & Millar, 1996; Millar & Lubben, 1996; Pillay et al., 2008; Rollnick et al., 2001; Sere et al., 2001, 1993) or individual data points. (Allie et al., 1998; Buffler et al., 2001; Davidowitz et al., 2001; Deardorff, 2001; Heinicke, 2012; Ibrahim, 2006; Rollnick et al., 2001; Sere et al., 2001; Volkwyn et al., 2008) This suggests that the correct answer on the original question with identical means may have come about for some reason other than deep conceptual understanding about spread. Students who do consider data spread in deciding whether two sets of data agree with each other or not will often say that the data agrees if the size of the spread is the same whether there is overlap or not. (Allie et al., 1998; Buffler et al., 2001; Coelho & Sere, 1998; Deardorff, 2001; Kung, 2005) Finally, even when students do compare by looking at data overlap, many do not consider the extent of overlap (Cauzinille-Marmeche et al., 1985), for example, in comparison to data spread within a single set of data. (Masnick & Morris, 2002) Regarding *combining* two sets of data with different spreads, research has documented the attitude that 'an average of numbers, obtained with different uncertainties, would have no sense.' (Sere et al., 1993) We can interpret these naïve ideas using the ontologies framework. Although we cannot predict exactly what value a measurement will yield, the spread of values obtained allows us to be more confident that the next measurement will lie somewhere within that spread. In this sense, we can

predict the extent of the randomness inherent in measurement. It seems plausible that students (even those who make correct decisions regarding spread in some situations!) struggle with this conceptually because, if they are assigning measurement into the 'deterministic' ontology, then the idea of randomness inherent in the measurement is not sensible. Conversely, if they are assigning measurement into the 'random' ontology, then the benefits of data spread would not be recognisable.

More advanced statistics

As discussed above, many researchers have documented a sense of carrying out learned procedures without an understanding of the meaning behind them. (Deardorff, 2001; Evangelinos et al., 2002; Lippmann, 2003; Lubben & Millar, 1996; Millar et al., 1994; Sere et al., 1993; Volkwyn et al., 2008) This issue of lacking conceptual knowledge behind the procedural tools used is present with more advanced tools as well. (Abbott, 2003; Garratt et al., 2010; Holmes, 2011; Sere et al., 1993) For example, the mistake of trying to calculate standard deviation by putting N-1 in the numerator has been documented. This mistake suggests 'that the student did not understand that the formula expresses that increasing the number of measurements improves the precision.'(Sere et al., 1993) Abbott discussed students measuring the mass of a paperclip (approximately 0.3 g) on a scale with a 0.1 g resolution multiple times so as to calculate standard deviation, despite only recording identical values. (Abbott, 2003)

The confidence interval is the range that contains with a certain probability the value to be measured. The probability is defined with the corresponding confidence level. For example, for 100 tosses of a penny, the number of heads will lie within the confidence interval of 46–54 with a confidence level of 95%. Research has documented students having trouble understanding confidence intervals. (Garratt et al., 2010; Goedhart & Verdonk, 1991) For example, some students think that the real result MUST lie within one standard deviation of the mean value (in fact, for normal distributed values, there is only a 67% confidence level.) (Goedhart & Verdonk, 1991) This view that the standard deviation sets bounds within which the true value *must* lie is consistent with students assigning measurement into the 'deterministic' ontology. With such a viewpoint, the idea that the true value need not lie within those bounds is not sensible, because it implies randomness.

Standard deviation is one tool commonly used to evaluate whether a data point is an outlier or not. As students do not understand standard deviation, it is perhaps unsurprising that they are at a loss to say what to do with outliers. When asked what to do with an anomalous data point, research has documented students dogmatically reporting that 'all readings must be used' or that it makes up part of the data spread and so needs to be included. Many students who opt to exclude the outlier do not explicitly justify the decision in terms of spread, explaining merely that it should be excluded because it was probably due to a mistake the experimenter made. (Allie et al., 1998; Rollnick et al., 2001) If students are not explicitly asked about outliers, many do not seem to notice them at all in data sets. (Davidowitz et al., 2001; Deardorff, 2001; Lubben & Millar, 1994; Millar & Lubben, 1996; Sere et al., 1993)

In making a line of best fit through data free-handed, research has documented students who either connect the dots or draw a line such that it passes perfectly through the maximum number of points. (*e.g.* Davidowitz et al., 2001; Lubben et al., 2001),

thinking, for example, that the dots not on the line must have been some mistake. (Rollnick et al., 2001) We can understand this behaviour in terms of assigning the relationship between two variables into the 'deterministic' ontology. Namely, we can imagine an argument of 'if we can predict how y depends upon x (it is a linear relationship), then measurements will predictably lie along that line. If they do not, it is because the experimenter needs more practice.' When using an objective method, like least mean squares linear regression (LMSLR), research has found students thinking that the process improves the accuracy, getting closer to the true value of the line. (Garratt et al., 2010) Garratt et al. argue that "The widely held belief that LMSLR increases the accuracy of (or decreases the error in) the final result almost certainly arises from the conventional use of the phrase 'line of best fit'. For a student drilled to appreciate the importance of accuracy, it will seem natural to associate 'best' with 'most accurate'." (Garratt et al., 2010)

Other naïve ideas described in the measurement literature pertaining to probability

Although conceptual understanding of the topic under investigation is important for deciding which data is meaningful to take (*e.g.* Millar et al., 1996), research has found students who evaluate results based entirely upon what they expect the result to be apriori rather than by being informed by the data. (Almekinders et al., 1998; Austin et al., 1991; Cauzinille-Marmeche et al., 1985; Coelho & Sere, 1998; Germann & Aram, 1996; Ibrahim, 2006; Lubben & Millar, 1994, 1996; Rollnick et al., 2001) For example, when given a spread of time data points for the time necessary for a tractor to travel a certain distance and asked to report one value, the highest value (9.8 seconds) was chosen with the explanation 'a tractor has no need to go fast. It isn't designed to have a high speed.' (Lubben & Millar, 1996) Research has also observed students doing the opposite, ignoring theory and focusing exclusively on data. For example, when presented with two potential curves to fit a set of data, some students suggested that computer use in handling the data would determine which curve is better (as opposed to considering the theoretical implications behind each curve).(Leach et al., 1998) The ontologies framework carries explanatory power here. It may be that students who ignore the data (because 'a tractor has no need to go fast', *e.g.*) are assigning measurement, or at least measurement of everyday life experiences, into the 'deterministic' ontology. Since tractors are not fast and measurement is a predictable process, these students might argue, it must be that measurements will reflect the slowness of the tractor. Similarly, students ignoring theory might be assigning measurement into the 'random' ontology. If measurement taking is a random process, these students might argue, then it follows no laws, and hence theory is not helpful in analysing it.

Austin et al. found that 13-year olds 'impose patterns on data when the data appear ambiguous, or when uncertainty surrounds a conclusion to be drawn from data' (Austin et al., 1991), and others have similarly found students prematurely making conclusions based on insignificant differences between samples. (Lubben & Millar, 1994; Millar et al., 1996) Researchers have also documented the opposite extreme in student thinking, that although one can talk about the behaviour of the particular sample (e.g. of shrimp), that these observations of the sample yield no power in making claims about the population. (Lubben & Millar, 1994) Almekinders et al. wrote 'students seem to lack the procedural understanding to decide that a large majority is a convincing basis for concluding that the experiment is effective.' (Almekinders et al., 1998) A proponent of the ontologies

framework might suggest that students prematurely imposing patterns on data reflects their assignment of measurement into the 'deterministic' ontology. Since measuring predictably yields information about the object of interest, these students might argue, it is only natural to assume that conclusions can be drawn from data. Similarly, a student who assigns measurement into the 'random' ontology might be expected to not extrapolate findings to reach a generalisation with the argument that if measurements are random, then there is no reason to expect that one can predict the result of the next measurement.

So far we have discussed the bulk of published student naïve ideas regarding Measurements and Uncertainties that relate to probability. We have argued that these naïve ideas can plausibly be connected in some way to assignment of the measurement process into either the 'deterministic' ontology or the 'random' ontology. However, not all of the probability-related naïve ideas described in the literature seem to connect to this underlying difficulty. This point will be revisited in the Limitations section below.

Discussion

The naïve ideas described in this paper are collected from myriad topics of physics and, within a given topic, are often considered to be 'misconceptions of (measurement, thermodynamics, etc.).' However, our decision to gather them together into one paper indicates our commitment to highlighting a shared theme. Namely, all of these naïve ideas seemingly relate in some way to probability, and we have considered the possibility that the naïve ideas in one topic and the naïve ideas in another topic are, in fact, the result of a more general probability-related student difficulty. Specifically, we have argued that the naïve ideas discussed in this literature review are caused, at least in part, by the puzzle explicated by mathematics education researchers: *how can a situation be random and predictable at the same time?* (Buechter et al., 2005; Gougis et al., 2017; Stavrou et al., 2003) Table 2 organises a few of the naïve ideas discussed in this literature review.[6] Some ideas, such as the idea that hot particles give their heat to cold particles (Kesidou & Duit, 1993), clearly illustrate our argument. For other naïve ideas, such as that condensation occurs

Table 2. Several examples of naïve ideas described in this literature review. The middle (right) column contains ideas we hypothesise to be caused, at least in part, by students categorising the relevant process as predictable (random).

Topic	Predictable	Random
Non-linear Systems	A chaotic pendulum changes which magnet it stops over because the Earth rotates (Duit et al., 1998)	There will be no similarity in dendprite appearance on electrodes from trial to trial (Komorek et al., 1999, Komorek et al., 2001)
Cosmology	Nothing existed before the Earth (Marques & Thompson, 1997)	None reported so far
Nuclear and Atomic Physics	A radioactive sample is safe after one half-life (Lijnse et al., 1990)	Coulomb's law is not relevant in thinking about ionisation energy (Taber, 1998)
Thermal Physics	Heat is transferred from one particle to another in the process of heat conduction (Kesidou & Duit, 1993)	Tiny droplets on the inside of a sealed bottle are due to a pressure difference (Chang, 1999) (see description)
Measurement	A single measurement is sufficient (*e.g.*; Allie et al., 1998)	Observations of the sample yield no power in making claims about the population (Lubben & Millar, 1994)

due to a pressure difference (Chang, 1999), our argument is more tenuous. With naïve ideas such as the former, we easily envision students attributing a characteristic of the collective to the individuals comprising that collective. Namely, recognising that heat is predictably transferred from the hot object to the cold object, students categorise the process as being predictable, and assume that the individual particles similarly behave predictably. With naïve ideas such as the latter regarding condensation, it is less clear what, and to what extent, the student was thinking when making the statement. It appears to be an issue of a failed memorisation attempt. We can envision, however, a student learning about random molecular motion and categorising the phenomena as random. Such a student would then be confused when learning that a temperature gradient predictably leads to changes in the random distribution of water molecules, because of the incompatibility perceived between randomness and predictions. In this state of confusion, the student may feel little choice but to attempt rote memorisation.

We have argued that many probability-related naïve ideas across physics topics originate, to varying degree, from the stance that 'random is incompatible with predictable'. In doing so, we have argued, more generally, that some of the naïve ideas discussed in, say, Quantum Physics, share a common cause as some of those in Measurements and Uncertainties. This argument of a 'common cause' was noticeably lacking in all of the articles we surveyed. In fact, very few of the sources (9 out of 70) discussed student reasoning outside of their own physics topic at all, and then only briefly. For example, 3 of these 9 articles aimed to see if the findings of Allie et al. regarding measurement difficulties of students in physics class also applied to students in chemistry (Davidowitz et al., 2001; Rollnick et al., 2001) and biology. (Almekinders et al., 1998) Lippmann's PhD Thesis considers the conditions under which the naïve ideas described above in Measurements and Uncertainties could be detrimental or beneficial in various job or everyday life-related situations. (Lippmann, 2003) De Posada mentioned how the topic of electrical conductivity is a bridge between physics and chemistry (De Posada, 1997) and Rozier and Viennot discussed how students learn to distinguish a particulate model from a continuous one in both fluid dynamics and quantum mechanics. (Rozier & Viennot, 1991) Harrison and Treagust, who discussed the naïve idea that the nucleus of an atom contains information for the atom, argued that the confusion arises in part because the same word is used in biology as in physics. (Harrison & Treagust, 1996) Taber, in discussing student failure to apply Coulomb's law in reasoning about the attraction of electrons to the nucleus, made it clear that students are not transferring physics knowledge into their chemistry lesson. (Taber, 1998) Finally, Kesidou and Duit looked at student understanding of conservation of energy in both thermodynamics and mechanics (in having students reason about a swinging pendulum). (Kesidou & Duit, 1993) None of these articles explicated the argument that naïve ideas in different probability-related fields may originate from a common cause, as we have argued. This is the primary contribution of our literature review.

This question of whether the diverse naïve ideas above stem from a common difficulty or not is a relevant question for educators, as the answer informs instruction. For example, it may prove to be more effective to treat that underlying difficulty than a host of individual naïve ideas. Proper treatment, however, requires proper diagnosis as to the specific nature of this underlying difficulty.

According to the ontologies view, the approach to take in helping students successfully overcome resilient misconceptions is to first teach them the correct ontology if it is not already present, and to then help them build conceptual understanding of the topic from within that ontology. (Slotta & Chi, 2006) 'Once students have successfully built such an alternative schema with its distinct set of properties, they can begin to assimilate new instruction (for example, about heat transfer) into the category.' (Chi, 2013)

In the case of teaching half-life, for example, it would follow that teachers should not try to utilise the existing ideas of learners if they are tied to a conception of radioactive decay classified using the 'deterministic' ontology. Rather, teachers should first teach an ontology where things exhibit *both* elements of randomness *and* predictability, and they should thereupon teach the half-life concept afresh in terms of that correct ontology.

Throughout this paper, we have utilised the ontologies framework as an example. However, this framework is not the only possibility for characterising the stance that 'random is incompatible with predictable'. For example, whereas the ontology framework conceptualises student understanding as beginning with categorising the relevant concepts ('water is a type of entity, not a type of process', e.g.), an alternative model is to instead conceptualise student understanding as an interaction of various cognitive elements. (e.g. (diSessa, 1993; DiSessa et al., 2016; Gupta et al., 2014; Hammer et al., 2005) With this model, we would conceptualise a student's view, for example, that radioactive decay is a predictable process, not as being the result of categorising within an intact ontology (the 'deterministic' ontology), but rather as being the result of various knowledge pieces coming together in a sort of constellation in that moment. DiSessa (DiSessa, 2017) describes such 'knowledge in pieces' frameworks in comparison to the ontologies framework as well the Theory Theory, which considers student reasoning to be theory-like.[7] (e.g. Carey, 2009; Lattery, 2016; Vosniadou & Skopeliti, 2014) A salient difference is that a knowledge in pieces framework attributes fluidity to student reasoning that is not easily accounted for in the other frameworks. Namely, it allows for the possibility that the stance 'random is incompatible with predictable' is not a rigid view, but rather one that shifts as different contexts grant students access to some knowledge pieces while prohibiting access to others. If student views regarding this stance do indeed prove to be more context-dependent, effective instructional intervention would look quite different from that proposed by Chi et al. above. As Hammer states,

> From one perspective, a teacher sees conceptions inherently inconsistent with expert knowledge; from the other, a teacher sees … knowledge elements that could contribute to expert understanding. The principal practical significance for a teacher is that the former implies the tasks of dismantling and replacing prior knowledge, whereas the latter suggests the task of modifying the organization and use of prior knowledge. (Hammer, 1996)

Limitations and conclusion

In this literature review, we examined naïve ideas from disparate topics that a Delphi study identified as having a common trait of being intertwined with probability. In total, there were seven physics topics we investigated: Non-linear Systems; Thermal Physics; Atomic, Nuclear and Particle Physics; Elementary Particle Physics; Quantum Physics; Cosmology; and Measurements and Uncertainties. Where possible, we utilised Ideas

Previas (Flores et al., 2014), an online database of naïve ideas. Namely, this was done for Cosmology (taken from the ideas collected under the heading 'Astonomy'), Atomic, Nuclear and Particle Physics; Elementary Particle Physics; and Quantum Physics (grouped together under 'Structure of Matter') and Thermal Physics. The foundation of Ideas Previas is Duit's database of articles documenting naïve ideas in various subjects (Duit, 2009), and we used this database for gathering naïve ideas regarding Non-linear Systems. Since even the Duit database did not contain articles relating to Measurements and Uncertainties, we carried out a more extensive and time-demanding process for this final topic. An important limitation of this paper is its scope. First, our scope was limited by our decision of which topics to investigate. This was informed by a Delphi study of experts, but there are undeniably naïve ideas outside of these topics that would fit well in this review. For example, the naïve idea discussed above that heat is transferred by individual hot molecules giving heat to individual cold molecules (Kesidou & Duit, 1993) is similar to the naïve idea we have frequently observed (and that is supported by some textbooks) that individual electrons carry energy from the battery to the bulb in an electric circuit. However, since this naïve idea is found within the topic of electromagnetism (not one of the topics we investigated), it is not included in our review.

Even within the topics that we did investigate, it is undeniably the case that there exist probability-related naïve ideas in, say, Thermal Physics that are not included in the Ideas Previas website and, consequently, not included in this literature review. This is particularly relevant since Ideas Previas only catalogues naïve ideas prior to 2008. Even prior to 2008 in these topics, we know of probability-related naïve ideas that Ideas Previas did not contain. In Quantum Physics, for example, Bethge (1988), Müller (2003), Petri (1996), and Wiesner (1996) observed the student argument that the position of the electron in a shell is expressed via a probability distribution only because the initial conditions are not adequately known. Correspondingly, students conflate probability and uncertainty. Furthermore, Bethge observed a persistent desire of students to attribute causality to quantum phenomena. In looking at single electrons travelling to a screen through a double-slit, for example, it is common for students to ask 'how does the electron know where it is supposed to go?' (Bethge, 1992) These and other naïve ideas, despite relating to probability, were missed by our methodology.

An additional limitation is the strength of the argument we can make. Namely, we certainly do not argue that *all* naïve ideas are due, even in part, to the stance that 'random is incompatible with predictable', nor do we argue that all naïve ideas in the topics we surveyed have this as a cause. We intentionally removed more than half of the naïve ideas from the 'Structure of Matter' and 'Thermodynamics' lists in Ideas Previas, for example, because they did not relate to probability. Although almost all of the probability-related naïve ideas we surveyed can be conceptualised as originating, at least to some degree, from the stance that 'random is incompatible with predictable', the argument remains only speculative for some of these ideas (for example, that condensation on the inside of a sealed bottle with a bit of water inside is due to a pressure difference (Chang, 1999)). Furthermore, in the topic Measurements and Uncertainties, there were two probability-related naïve ideas which seem disconnected to the stance that 'random is incompatible with predictable'. Specifically, Sere et al. documented a student measuring the diameter of a wire with the idea that an arithmetic mean of multiple measurements of a wire's diameter does not yield the diameter of the actual wire being measured, but rather the diameter of a

perfectly-cylindrical theoretical wire. (Sere et al., 1993) Furthermore, Lippmann found that students think that, so long as something is common between trials, then it need not be considered. For example, since background radiation is present for each trial and is the same, you don't need to subtract it when calculating the half-life of a radioactive sample. (Lippmann, 2003)

Despite these limitations, we feel that this review is sufficiently comprehensive to be able to make the argument that many naïve ideas from disparate topics stem, to varying degree, from a common student difficulty with probability (namely, the stance that 'random is incompatible with predictable'). Despite this plausible commonality, very few of the articles we reviewed discussed student reasoning in topics other than the one of focus. As an example of what analysis of this underlying difficulty could look like, we considered how one might apply the ontologies framework to inspect the various naïve ideas. Although the ontologies framework attributes stability to student reasoning, there exists ample research in physics education to suggest that student cognition is often more fluid in nature (*e.g.* diSessa, 1993; Hammer, 1996; Hammer et al., 2005). It is thus plausible that the 'random is incompatible with predictable' stance (like naïve ideas themselves) need not take the form of two mutually-exclusive and intact ontologies ('arbitrary' and 'random'), but could rather be a collection of smaller knowledge pieces that is only assembled into this stance under certain contexts. However, few of the articles reviewed in this paper considered context-dependency of student reasoning. The topic with the largest percentage of papers that made such considerations was Measurements and Uncertainties, and then only 9 of the sources reviewed did so (Abbott, 2003; Allie et al., 1998; Heinicke, 2012; Kung, 2005; Leach et al., 1998; Lippmann, 2003; Lubben et al., 2001; Sere et al., 2001; Volkwyn et al., 2008). One of these 9 sources, Lippmann's PhD thesis, criticises the perspective of Buffler et al. (2001) that the naïve ideas we described above (many of which are collectively referred to as the 'point paradigm') are, implicitly, misconceptions that instructors should aim to help students transcend (to arrive at the 'set paradigm'). Lippmann writes

> Underlying Buffler et al.'s research is a misconceptions perspective. They wish to stop students from reasoning from the point paradigm, and teach them to reason only from the set paradigm. We would like to propose that a [knowledge-in-pieces] framework is more suitable. Students do not need to throw out the point paradigm and replace it with the set paradigm; instead, they need the skills to use either paradigm and to determine which paradigm to use. Different types of reasoning are productive depending on the type of question. (p. 28 of Lippmann, 2003)

In this paper, we have speculated that what appears at first to be a host of disparate naïve ideas is in fact driven by a common stance of 'random is incompatible with predictable'. We can similarly only speculate whether this stance is rigidly context-independent (in which case the ontologies framework seems well suited to describe the situation) or if, given a different phrasing of the question or a more familiar context, a student might actually shift to the correct view. Although we have not gone beyond speculation here, we think that this latter point in particular is a research question that deserves attention, and we intend to investigate it in the near future.

Notes

1. In fact, the question of whether things are truly random or not is a philosophical one that has not been resolved. (Heuer, 2013; Schurz, 2015) The important point here is that the students, in their argument that they must gather more information to predict the trajectory of the ball, are blinded to a far more practical approach afforded by probability.
2. Chi et al. treat the term 'misconception' as an intact knowledge structure that interferes with student learning. We use the term 'naïve idea' to reflect agnosticism regarding how 'intact' these knowledge structures are. We discuss this more in the Discussion, where we mention frameworks that allow for fluidity in student reasoning.
3. It is possible that our description of the 'random' ontology here is different than what Chi et al. had intended, as they do not describe the ontology. Regardless, this does not affect our claim that naïve ideas across a variety of physics topics are plausibly enabled by students assigning phenomena into one of two mutually exclusive ontologies.
4. The fact that so few naïve ideas have been documented regarding student understanding of cosmology is an interesting side finding of our research, and it demonstrates room for subsequent research into this field.
5. The Koch curve is formed from a line segment by repeating the following three rules: 1) divide the line segment into three segments of equal length, 2) draw an equilateral triangle that has the middle segment from step 1 as its base and points outward, 3) remove the line segment that is the base of the triangle from step 2.
6. As discussed in the introduction, mathematics education researchers have found that many students reject ideas of randomness. (Buechter et al., 2005; Green, 1983; Kahneman & Tversky, 1973) Correspondingly, one would expect a student to categorise most things in science as predictable, as reflected in this table and throughout this review.
7. Whereas DiSessa writes from a Pieces perspective, Vosniadou and Skopeliti compare the three frameworks from the Theory perspective. (Vosniadou & Skopeliti, 2014) See also (Vosniadou, 2013) for thorough and recent descriptions of the three frameworks.

Disclosure statement

We had no financial interest or benefit that has arisen from the direct applications of our research.

Notes on contributors

Michael M. Hull is a postdoctoral research associate focusing on student understanding of radioactivity.

Alexandra Jansky is a postdoctoral fellow specializing in student understanding of randomness and radiation detection.

Martin Hopf is the head of the Austrian Educational Competence Centre, which investigates student understanding in a range of topics ranging from mechanics to radiation and designs curriculum responsively.

ORCID

Michael M. Hull 🆔 http://orcid.org/0000-0003-1551-9553
Alexandra Jansky 🆔 http://orcid.org/0000-0003-0594-2390

References

Abbott, D. S. (2003). *Assessing student understanding of measurement and uncertainty*. North Carolina State University.

Allie, S., Buffler, A., Kaunda, L., Campbell, B., & Lubben, F. (1998). First-year physics students' perceptions of the quality of experimental measurements. *International Journal of Science Education*, *20*(4), 447–459. https://doi.org/10.1080/0950069980200405

Almekinders, R., Thijs, G., & Fred, L. (1998). Development of procedural understanding among South African science students at pre-tertiary education level. *Journal of Biological Education*, *33*(1), 33–38. https://doi.org/10.1080/00219266.1998.9655634

Alsop, S. (2001). Living with and learning about radioactivity: A comparative conceptual study. *International Journal of Science Education*, *23*(3), 263–281. https://doi.org/10.1080/095006901750066510

Austin, R., Holding, B., Bell, J., & Daniels, S. (1991). Patterns and relationships in school science. *Assess-Ment Matters*, *7(1)*.

Barragués, J. I., Guisasola, J., & Morais, A. (2006). Chance and probability: What do they mean to university engineering students? *International Journal of Mathematical Education in Science and Technology*, *37*(8), 883–900. https://doi.org/10.1080/00207390600818948

Batanero, C., Chernoff, E. J., Engel, J., Lee, H. S., & Sánchez, E. (2016). Research on teaching and learning probability. In *Research on teaching and learning probability* (pp. 1–33). Springer. https://doi.org/10.1007/978-3-319-31625-3_1

Bell, T., & Euler, M. (2001). Learning about complex systems: How much instruction do we need to foster autonomous model construction. In D. Psillos, P. Kariotoglou, V. Tselfes, G. Bisdikian, G. Fassoulopoulos, E. Hatzikraniotis, & M. Kallery (Eds.), *Proceedings of the Third International Conference on Science Education Research in the Knowledge Based Society, Vol. 1*. (pp. 291–293). Thessaloniki, Greece.

Bell, T. (2007). Konzeptentwicklung in einer Lernprozessstudie im Breich "komplexe Systeme und Selbstregulation.". *Physik Und Didaktik in Schule Und Hochschule*, *2*(6), 59–71.

Bethge, T. (1988). *Aspekte des Schülervorverständnisses zu grundlegenden Begriffen der Atomphysik*. Universität Bremen.

Bethge, T. (1992). Vorstellung von Schülerinnen und Schülern zu Begriffen der Atomphysik. In H. F. Hrsg (Ed.), *Quantenphysik in der Schule* (pp. 215–233). IPN.

Boeing, G. (2016). Visual analysis of nonlinear dynamical systems: Chaos, fractals, self-similarity and the limits of prediction. *Systems*, *4*(4), 37. https://doi.org/10.3390/systems4040037

Bowen, A., Reid, D., & Koretsky, M. (2015). Development of interactive virtual laboratories to help students learn difficult concepts in thermodynamics. *Chemical Engineering Education*, *49*(4), 229–238.

Buechter, A., Hussmann, S., Leuders, T., & Prediger, S. (2005). Den Zufall im Griff?–Stochastische Vorstellungen fördern. *Praxis Der Mathematik in Der Schule*, *47*(4), 1–7.

Buecker, N., Komorek, M., & Duit, R. (1999). Experimente, elementarisierungen und schuelervorstellungen im bereich fraktales wachstum. In: R. Brechel (Ed.): Zur Didaktik der Physik und Chemie, Probleme und Perspektiven - Vorträge zur GDCP-Jahrestagung in Essen 1998, Alsbach: Leuchtturm, 229–231

Buffler, A., Allie, S., & Lubben, F. (2001). The development of first year physics students' ideas about measurement in terms of point and set paradigms. *International Journal of Science Education*, *23*(11), 1137–1156. https://doi.org/10.1080/09500690110039567

Carey, S. (2009). *The origin of concepts*. Oxford University Press.

Cauzinille-Marmeche, E., Meheut, M., Sere, M., & Weil-Barais, A. (1985). The influence of a priori ideas on the experimental approach. *Science Education*, *69*(2), 201–211. https://doi.org/10.1002/sce.3730690212

Chang, J.-Y. (1999). Teachers college students' conceptions about evaporation, condensation, and boiling. *Science Education*, *83*(5), 511–526. https://doi.org/10.1002/(SICI)1098-237X(199909)83:5<511::AID-SCE1>3.0.CO;2-E

Chi, M. T. H. (2013). Two kinds and four sub-types of misconceived knowledge, ways to change it, and the learning outcomes. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (2nd ed.). Routledge. pp 61-82.

Chi, M. T. H., & Slotta, J. D. (1993). The ontological coherence of intuitive physics. *Cognition and Instruction*, *10*(2–3), 249–260. https://doi.org/10.1080/07370008.1985.9649011

Chi, M. T. H., Slotta, J. D., & De Leeuw, N. (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction*, *4* (1), 27–43. https://www.sciencedirect.com/science/article/pii/0959475294900175.

Coelho, S. M., & Sere, M. (1998). Pupils' reasoning and practice during hands-on activities in the measurement phase. *Research in Science & Technological Education*, *16*(1), 79–96. https://doi.org/10.1080/0263514980160107

Cros, D., Chastrette, M., & Fayol, M. (1988). Conceptions of second year university students of some fundamental notions in chemistry. *International Journal of Science Education*, *10*(3), 331–336. https://doi.org/10.1080/0950069880100308

Davidowitz, B., Lubben, F., & Rollnick, M. (2001). Undergraduate science and engineering students' understanding of the reliability of chemical data. *Journal of Chemical Education*, *78*(2), 247. https://doi.org/10.1021/ed078p247

De Posada, J. M. (1997). Conceptions of high school students concerning the internal structure of metals and their electric conduction: Structure and evolution. *Science Education*, *81*(4), 445–467. https://doi.org/10.1002/(SICI)1098-237X(199707)81:4<445::AID-SCE5>3.0.CO;2-C

Deardorff, D. L. (2001). *Introductory physics students' treatment of measurement uncertainty*. North Carolina State University.

Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M., Wales, P. W., (2014). No TitleDefining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology*, *67*(4), 401–409. https://doi.org/10.1016/j.jclinepi.2013.12.002

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, *10*(2–3), 105–225. https://doi.org/10.1080/07370008.1985.9649008

DiSessa, A. A., Sherin, B., & Levin, M. (2016). Knowledge analysis: An introduction. In A. DiSessa, M. Levin, & N. Brown (Eds.), *Knowledge and interaction: A synthetic agenda for the learning sciences* (pp. 30–71). Routledge.

DiSessa, A. A. (2017). Conceptual change in a microcosm: Comparative analysis of a learning event. *Human Development*, *60*(1), 1–37. https://doi.org/10.1159/000469693

Döhrmann, M. (2005). Schülervorstellung zum Begriff "Zufall.". In In: Graumann, G. (Ed.): Beitraege zum Mathematikunterricht 2005 – Vortraege auf der 39. Tagung für Didaktik der Mathematik, Bielefeld. Franzbecker: Hildesheim, S. 167-170 .

Duit, R. (2009). *Bibliography - STCSE students' and teachers' conceptions and science education*. Retrieved July 31, 2019, from http://archiv.ipn.uni-kiel.de/stcse/

Duit, R., Roth, W.-M., Komorek, M., & Wilbers, J. (1998). Conceptual change cum discourse analysis to understand cognition in a unit on chaotic systems: Towards an integrative perspective on learning in science. *International Journal of Science Education*, *20*(8), 1059–1074. https://doi.org/10.1080/0950069980200904

Erickson, G. L. (1979). Children's conceptions of heat and temperature. *Science Education*, *63*(2), 221–230. https://doi.org/10.1002/sce.3730630210

Evangelinos, D., Psillos, D., & Valassiades, O. (2002). An investigation of teaching and learning about measurement data and their treatment in the introductory physics laboratory. In *Teaching and learning in the science laboratory, Psillos, Dimitris, Niedderer, H. (eds.)* (pp. 179–190). Springer.

Fischler, H., & Lichtfeldt, M. (1992). Modern physics and students' conceptions. *International Journal of Science Education*, *14*(2), 181–190. https://doi.org/10.1080/0950069920140206

Flores, C. F., Vega, M. E. J., Tovar, M. M. E., & Bello, G. S. (2014). *Ideas Previas*. Retrieved July 31, 2019, from http://www.ideasprevias.ccadet.unam.mx:8080/ideasprevias/searching.htm

Garratt, J., Horn, A., & Tomlinson, J. (2010). Misconceptions about Error. *University Chemistry Education*, *4(2)*, 2.

Germann, P., & Aram, R. (1996). Student performances on the science processes of recording data, analyzing data, drawing conclusions, and providing evidence. *Journal of Research in Science Teaching*, *33*(7), 773–798. https://doi.org/10.1002/(SICI)1098-2736(199609)33:7<773::AID-TEA5>3.0.CO;2-K

Goedhart, M. J., & Verdonk, A. H. (1991). The development of statistical concepts in a design-oriented laboratory course in scientific measuring. *Journal of Chemical Education*, *68*(12), 1005. https://doi.org/10.1021/ed068p1005

Gott, R., & Duggan, S. (1995). *Investigative work in the science curriculum. Developing science and technology education*. Open University Press.

Gougis, R. D., Stomberg, J. F., O'Hare, A. T., O'Reilly, C. M., Bader, N. E., Meixner, T., & Carey, C. C. (2017). Post-secondary science students' explanations of randomness and variation and implications for science learning. *International Journal of Science and Mathematics Education*, *15*(6), 1039–1056. https://doi.org/10.1007/s10763-016-9737-7

Green, D. R. (1983). School pupils' probability concepts. *Teaching Statistics*, *5*(2), 34–42. https://doi.org/10.1111/j.1467-9639.1983.tb00493.x

Griffiths, A. K., & Preston, K. R. (1992). Grade-12 students' misconceptions relating to fundamental characteristics of atoms and molecules. *Journal of Research in Science Teaching*, *29*(6), 611–628. https://doi.org/10.1002/tea.3660290609

Gupta, A., Elby, A., & Conlin, L. (2014). How substance-based ontologies for gravity can be productive: A case study. *Physical Review Special Topics - Physics Education Research*, *10*(1), 010113. https://doi.org/10.1103/PhysRevSTPER.10.010113

Hackling, M. W., & Garnett, P. J. (1992). Expert — Novice differences in science investigation skills. *Research in Science Education*, *22*(1), 170–177. https://doi.org/10.1007/BF02356893

Häder, M. (2009). *Delphi-Befragungen: Ein Arbeitsbuch*. Springer-Verlag.

Hammer, D. (1996). Misconceptions or p-prims: How may alternative perspectives of cognitive structure influence instructional perceptions and intentions. *Journal of the Learning Sciences*, *5* (2), 97–127. https://doi.org/10.1207/s15327809jls0502_1

Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. *Transfer of Learning from a Modern Multidisciplinary Perspective*, *89*. http://umdperg.pbworks.com/w/file/fetch/51074580/Transfer_chapter_final.pdf. 89-121

Harrison, A. G., & Treagust, D. F. (1996). Secondary students' mental models of atoms and molecules: Implications for teaching chemistry. *Science Education*, *80*(5), 509–534. https://doi.org/10.1002/(SICI)1098-237X(199609)80:5<509::AID-SCE2>3.0.CO;2-F

Heinicke, S. (2012). *Aus Fehlern Wird Man Klug: Eine Genetisch-Didaktische Rekonstruktion des Messfehlers*. Logos Verlag Berlin GmbH. https://books.google.com/books?hl=en&id=pPNI9uKfCpIC&oi=fnd&pg=PA1&dq=Susanne+Heinicke+%22Aus+Fehlern+wird+man+klug,+Eine+Genetisch-Didaktische+Rekonstruktion+des+Messfehlers%22&ots=-CcYx-NYXl&sig=8CR3X6xJeorsxl5kAR2a0bPrR_0

Heuer, P. (2013). Gibt es Zufälle? In S. Rödl, & H. Tegtmeyer (Ed.), *Sinnkritisches Philosophieren. Berlin: DeGruyter.* (pp. 311–336).

Hewson, M. G. A. B., & Hamlyn, D. (1984). The influence of intellectual environment on conceptions of heat. *European Journal of Science Education*, *6*(3), 245–262. https://doi.org/10.1080/0140528840060306

Hewson, Mariana. A'B., and Daryl Hamlyn. "the influence of intellectual environment on conceptions of heat." European Journal of Science Education 6.*3* (1984): 245–262

Hogan, K., & Fisherkeller, J. (1996). Representing students' thinking about nutrient cycling in ecosystems: Bidimensional coding of a complex topic. *Journal of Research in Science Teaching*, *33*(9), 941–970. https://doi.org/10.1002/(SICI)1098-2736(199611)33:9<941::AID-TEA1>3.0.CO;2-V

Holmes, N. G. (2011). *The invention support environment: Using metacognitive scaffolding and interactive learning environments to improve learning from invention*. The University Of British Columbia.

Hwang, B. T., & Hwang, H. W. (1990). A study of cognitive development of the concepts of solution. (CITED BY chang, Jin-Yi. (1999). Teacher's college students' conceptions about evaporation, condensation and boiling. *Science Education*, *83* (5), 511–526,p. 513. Paipei

IBO. (2014). *Diploma programme: Physics guide, For first examination in 2016. International Baccalaureate Organization*. Geneva. https://ibphysics.org/wp-content/uploads/2016/01/ib-physics-syllabus.pdf

Ibrahim, B. B. (2006). *The relationship between views of the nature of science and views of the nature of scientific measurement*. University of Cape Town.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*(4), 237–251. https://doi.org/10.1037/h0034747

Keil, F. C. (1979). *Semantic and conceptual development: An ontological perspective*. Harvard University Press.

Kesidou, S., & Duit, R. (1993). Students' conceptions of the second law of thermodynamics—an interpretive study. *Journal of Research in Science Teaching*, *30*(1), 85–106. https://doi.org/10.1002/tea.3660300107

Komorek, M. (1998). Lernprozessstudie zum deterministischen Chaos. *Zeitschrift Fuer Didaktik Der Naturwissenschaften*, *5*(3), 3–22.

Komorek, M., Duit, R., Naujack, B., & Buecker, N. (1999). Learning process studies in the field of fractals. In A. M. Komorek, H. Behrendt, H. Dahncke, R. Duit, & W., . K. Graeber (Eds.), *Research in science education - past, present, and future* (Vol. 1, pp. 131–133). IPN Kiel.

Komorek, M., Duit, R., Naujack, B., Buecker, N., & Naujak, B. (2001). Learning process studies in the field of fractals. In H. Behrendt, H. Dahncke, R. Duit, W. Graeber, M. Komorek, & K. A. Dordrecht (Eds.), *Research in science education - past, present, and future Vol.1* (pp. 95–100). Kluwer Academic Publishers.

Kung, R. L. (2005). Teaching the concepts of measurement: An example of a concept-based laboratory course. *American Journal of Physics*, *73*(8), 771. https://doi.org/10.1119/1.1881253

Lattery, M. J. (2016). *Deep learning in introductory physics: Exploratory studies of model-based reasoning*. IAP.

Leach, J., Millar, R., Ryder, J., Sere, M.-G., Hammelev, D., Niedderer, H., Leach Lis, J. (1998). *Labwork in science education survey 2: Students' images of science as they relate to labwork learning*. https://s3.amazonaws.com/academia.edu.documents/34112485/1998-LSE-WP4.pdf?AWSAccessKeyId=A K I A I W O W Y Y G Z 2 Y 5 3 U L 3 A & E x p i r e s = 1 5 2 8 9 0 2 6 3 1 & S i g n a t u r e = CPIBmLvVI5qsyOQtbDLQ0rER550%3D&response-content-disposition=inline%3Bfilename%3DSURVEY_2_STUDENTS_IMAGES_OF_SCIENCE

Lijnse, P. L., Eijkelhof, H. M. C., Klaassen, C. W. J. M., & Scholte, R. L. J. (1990). Pupils' and mass-media ideas about radioactivity. *International Journal of Science Education*, *12*(1), 67–78. https://doi.org/10.1080/0950069900120106

Lippmann, R. F. (2003). *Students' understanding of measurement and uncertainty in the physics laboratory: Social construction, underlying concepts, and quantitative analysis. ProQuest Dissertations And Theses; Thesis (Ph.D.)–University of Maryland, College Park, 2003.; Publication Number: AAI3112498; ISBN: 9780496600106; Source: Dissertation Abstracts International, Volume: 64-11, Section: B, page: 5560.*, 349. University of Maryland College Park. http://adsabs.harvard.edu/abs/2003PhDT

Lubben, F., Campbell, B., Buffler, A., & Allie, S. (2001). Point and set reasoning in practical science measurement by entering university freshmen. *Science Education*, *85*(4), 311–327. https://doi.org/10.1002/sce.1012

Lubben, F., & Millar, R. (1994). *A survey of the understanding oj children aged 11-16 oj key ldeas about evidence in science*. York.

Lubben, F., & Millar, R. (1996). Children's ideas about the reliability of experimental data. *International Journal of Science Education*, *18*(8), 955–968. https://doi.org/10.1080/0950069960180807

Manthei, U. (1980). Zur genetischen Begriffsdifferenzierung und Begriffspraezisierung. *Phsisik in Der Schule*, *18*(9), 389–398.

Marques, L., & Thompson, D. (1997). Portuguese students' understanding at ages 10-11 and 14-15 of the origin and nature of the earth and the development of life. *Research in Science & Technological Education*, *15*(1), 29–51. https://doi.org/10.1080/0263514970150103

Masnick, A. M., & Morris, B. J. (2002). *Reasoning from data: The effect of sample size and variability on children's and adults' conclusions*. Proceedings of the annual meeting of the cognitive science society.

McDermott, L. C., & Redish, E. F. (1999). Resource letter: PER-1: physics education research. *American Journal of Physics*, *67*(9), 755–767. https://doi.org/10.1119/1.19122

Millar, R., Gott, R., Lubben, F., & Duggan, S. (1996). Children's performance of investigative tasks in science: A framework for considering progression. In M. Hughes (Ed.), *Progression in learning* (pp. 82–108). Multilingual Matters.

Millar, R., & Gill, J. S. (1996). School students' understanding of processes involving radioactive substances and ionizing radiation. *Physics Education*, *31*(1), 27–33. https://doi.org/10.1088/0031-9120/31/1/019

Millar, R., & Lubben, F. (1996). Knowledge and action: Students' understanding of the procedures of scientific enquiry. In Geoff Welford, Jonathan Osborne, Phil Scott. *In research in science education in Europe* (pp. 166). Routledge.

Millar, R., Lubben, F., Gott, R., & Duggan, S. (1994). Investigating in the school science laboratory: Conceptual and procedural knowledge and their influence on performance. *Research Papers in Education*, *9*(2), 207–248. https://doi.org/10.1080/0267152940090205

Müller, R. (2003). *Quantenphysik in der Schule. Studien zum Physiklernen*. Logos.

Petri, J. (1996). *Der Lernpfad eines Schülers in der Atomphysik*. Mainz.

Pillay, S., Buffler, A., Lubben, F., & Allie, S. (2008). Effectiveness of a GUM-compliant course for teaching measurement in the introductory physics laboratory. *European Journal of Physics*, *29*(3), 647–659. https://doi.org/10.1088/0143-0807/29/3/024

Reiner, M., Slotta, J. D., Chi, M. T. H., & Resnick, L. B. (2000). Naive physics reasoning: A commitment to substance-based conceptions. *Cognition and Instruction*, *18*(1), 1–34. https://doi.org/10.1207/S1532690XCI1801_01

Rollnick, M., Dlamini, B., Lotz, S., & Lubben, F. (2001). Views of South African chemistry students in university bridging programs on the reliability of experimental data. *Research in Science Education*, *31*(4), 553–573. https://doi.org/10.1023/A:1013102108541

Roth, W.-M., & Duit, R. (2003). Emergence, flexibility, and stabilization of language in a physics classroom. *Journal of Research in Science Teaching*, *40*(9), 869–897. https://doi.org/10.1002/tea.10114

Rozier, S., & Viennot, L. (1991). Students' reasonings in thermodynamics. *International Journal of Science Education*, *13*(2), 159–170. https://doi.org/10.1080/0950069910130203

Saricayir, H., Ay, S., Comek, A., Cansiz, G., & Uce, M. (2016). Determining students' conceptual understanding level of thermodynamics. *Journal of Education and Training Studies*, *4*(6), 6. https://doi.org/10.11114/jets.v4i6.1421

Schurz, G. (2015). Probleme des objektiv-statistischen Wahrscheinlichkeitsbegriffs. In W. de Gruyter (Ed.), *Wahrscheinlichkeit*. Berlin/Boston. pp. 39-64.

Sere, M.-G., Fernandez-Gonzalez, M., Gallegos, J. A., Gonzalez-Garcia, F., De Manuel, E., Perales, F. J., & Leach, J. (2001). Images of science linked to labwork: A survey of secondary school and university students. *Research in Science Education*, *31*(4), 499–523. https://doi.org/10.1023/A:1013141706723

Sere, M.-G., Journeaux, R., & Larcher, C. (1993). Learning the statistical analysis of measurement errors. *International Journal of Science Education*, *15*(4), 427–438. https://doi.org/10.1080/0950069930150406

Sharp, J. G. (1996). Children's astronomical beliefs: A preliminary study of Year 6 children in southwest England. *International Journal of Science Education*, *18*(6), 685–712. https://doi.org/10.1080/0950069960180604

Shaughnessy, J. M., & Ciancetta, M. (2002). *Students' understanding of variability in a probability environment*. Proceedings of the Sixth International Conference on Teaching Statistics: Developing a Statistically Literate Society. http://iase-web.org/documents/papers/icots6/6a6_shau.pdf

Shaughnessy, J. M., & Zawojewski, J. S. (1999). Secondary students' performance on data and chance in the 1996 NAEP. *The Mathematics Teacher*, *92* (8), 713. http://search.proquest.com/openview/1ef2b981b447c8612498b0abbc049527/1?pq-origsite=gscholar&cbl=41299

Slotta, J. D., & Chi, M. T. H. (2006). Helping students understand challenging topics in science through ontology training. *Cognition and Instruction*, *24*(2), 261–289. https://doi.org/10.1207/s1532690xci2402_3

Slotta, J. D., Chi, M. T. H., & Joram, E. (1995). Assessing students' misclassifications of physics concepts: An ontological basis for conceptual change. *Cognition and Instruction*, *13*(3), 373–400. https://doi.org/10.1207/s1532690xci1303_2

Sriyansyah, S. P., Karim, S., Saepuzaman, D., & Suhandi, A. (2015). *Investigation of students' scientific consistency and learning difficulties in the first law of thermodynamics*. International conference on mathematics, science, and education.

Stavrou, D., & Duit, R. (2014). Teaching and learning the interplay between chance and determinism in nonlinear systems. *International Journal of Science Education*, *36*(3), 506–530. https://doi.org/10.1080/09500693.2013.802056

Stavrou, D., Duit, R., & Komorek, M. (2008). A teaching and learning sequence about the interplay of chance and determinism in nonlinear systems. *Physics Education*, *43*(4), 417–422. https://doi.org/10.1088/0031-9120/43/4/011

Stavrou, D., Komorek, M., & Duit, R. 2003. Schülervorstellungen über das Wechselspiel von Determinismus und Zufall. *Außerschulisches Lernen in Physik und Chemie Band 23* (Vol. 23299–301).A. Pitton. LIT Verlag.

Stavrou, D., Komorek, M., & Duit, R. (2005). Didaktische Rekonstruktion des Zusammenspiels von Zufall und Gesetzmäßigkeit in der nichtlinearen Dynamik. *Zeitschrift Für Didaktik Der Naturwissenschaften*, *11*, 147–164.

Stehlik, S., & Pade, J. (2006). Schüler begegnen der Nichtlinearen Dynamik - ein Projekt zur Untersuchung von Schülervorstellungen anhand des chaotischen Wasserrads. In *Frühjahrstagung des Fachverbands Didaktik der Physik der Deutschen Physikalischen Gesellschaft*. Lehmanns Media.

Strike, K. A., & Posner, G. J. (1982). Conceptual change and science teaching. *European Journal of Science Education*, *4*(3), 231–240. https://doi.org/10.1080/0140528820040302

Summers, M. K. (1983). Teaching heat–an analysis of misconceptions. *School Science Review*, *64* (229), 670–676. https://eric.ed.gov/?id=EJ283122

Taber, K. S. (1998). The sharing-out of nuclear attraction: Or 'I can't think about physics in chemistry'. *International Journal of Science Education*, *20*(8), 1001–1014. https://doi.org/10.1080/0950069980200807

Ure, M. C. D., & Colinvaux, D. (1989). Developing adults' views on the phenomenon of change of physical state in water. *International Journal of Science Education*, *11*(2), 153–160. https://doi.org/10.1080/0950069890110204

Vacc, N. N. (1999). Exploring fractal geometry with children. *School Science and Mathematics*, *99*(2), 77–83. https://doi.org/10.1111/j.1949-8594.1999.tb17451.x

Volkwyn, T. S., Allie, S., Buffler, A., & Lubben, F. (2008). Impact of a conventional introductory laboratory course on the understanding of measurement. *Physical Review Special Topics - Physics Education Research*, *4*(1), 010108. https://doi.org/10.1103/PhysRevSTPER.4.010108

Vosniadou, S. (Ed.). (2013). *International handbook of research on conceptual change* (2nd ed.). Routledge.

Vosniadou, S., & Skopeliti, I. (2014). Conceptual change from the framework theory side of the fence. *Science & Education*, *23*(7), 1427–1445. https://doi.org/10.1007/s11191-013-9640-3

Warwick, P., Linfield, R. S., & Stephenson, P. (1999). A comparison of primary school pupils' ability to express procedural understanding in science through speech and writing. *International Journal of Science Education*, *21*(8), 823–838. https://doi.org/10.1080/095006999290318

Wiesner, H. (1996). Verständnisse von Leistungskursschülern über Quantenphysik (1/2). Ergebnisse mündlicher Befragungen. In *Physik in der Schule* (pp. 34, 95–99, 136–140). Köln: Aulis Verlag.

Wilbers, J. D., & Duit, R. (2001). On the micro-structure of analogical reasoning: The case of understanding chaotic systems. In I. M. Komorek, H. Behrendt, H. Dahncke, R. Duit, W. Graeber, & K. A. Kiel (Eds.), *Research in science education - past, present, and future* (pp. 205–210). Kluwer Academic Publishers.

Appendix Naïve Ideas taken from Ideas Previas(Flores et al., 2014)

Table A1. Naïve ideas, printed verbatim, that are listed under the heading 'Thermodynamics' in Ideas Previas and that relate to probability.

Source	Naïve ideas
(Chang, 1999)	A tightly capped clear plastic bottle filled with a small amount of water is kept indoors for a few days. As one can see, many tiny water droplets appear on the inner surface of the bottle … 1) They come from the air that is within the bottle.
	… 2) They come from the difference of temperature between the outside and the inside of the bottle.
	… 3) They come from the pressure difference present at the outside and inside of the bottle.
	If one places a piece of glass over the top of a cup of just-boiled water, after a few minutes, tiny water droplets appear under the glass. … 4) They come from hot and cold air when they encounter.
	A bottle of liquid drink is taken out of the refrigerator for not too long. As one can see, there are many fine water droplets on the surface of the bottle. … 5) They come from cold water evaporated when encountering heat.
	… 6) They come from the cold air condensed when it encounters heat.
	… 7) They come from the meeting of cold and hot air.
	… 8) They are produced when cold drink encounters heat.
	In a room there is a clear plastic bottle filled with a small amount of water. Its lid is removed and the water is left undisturbed for several days. What will happen after several days? Why does such a phenomenon take place? … 9) After several days water vaporises when it changes to the gas state and combines with some moleculas in air to run upwards.
	… 10) Water will gradually diminish, because it will flow into the air.
(Ure & Colinvaux, 1989)	When a mirror is placed over a pot of boiling water, such water that appears on the mirror … 11) comes from the surrounding air.
	… 12) It comes from that water passes through the glass to the outside.
	… 13) It comes from the sweat, which is made of ice because the ice with the temperature, will tend to sweat.
(Erickson, 1979)	14) Everything contains air bubbles. Some of the bubbles might contain hot air and some cold air.
	15) Some things like ice and sugar cubes melt because they contain air bubbles.
	16) All objects contain a mixture of heat and cold.
	17) Hot substances contain fumes, and when they cool down these fumes escape gradually into the air.
(Hewson & Hamlyn, 1984)	18) Heat is movement of warm air from the fire.
	19) Heat energy makes the energy inside the particles increase and the particles expands.
	20) When heating steel and water, the molecules get bigger.
	21) Heated particles increase in number by splitting or multiplying so that there are more particles of the same substance (occupying more space).
	22) This heat it comes [into the ball] … is full of something … and when it gets there it's going to make the thing increase its size … and when it cools, [the heat] goes back with all its constituents and then [the ball] returns to its size.
	23) Heat is energy, which is like waves or rays.
	24) Heat travels faster in light materials because the particles are not so packed.
	25) When the rod is heated the flame destroys the inner composition.
(Hwang & Hwang, 1990)	26) When the water vapour is evaporating the solute particles have come out from the solution together
	with the solvent particles.
(Kesidou & Duit, 1993)	27) The heat is movement of particles.
	28) Temperature is a speed of particles.
	29) Particles expand when heated.
	30) Heat is transferred from one particle to another in the process of heat conduction.
(Manthei, 1980)	31) Particles rub against one another and hence become warmer.

(*Continued*)

Table A1. (Continued).

Source	Naïve ideas
(Rozier & Viennot, 1991)	32) An adiabatic compression of a perfect gas results in an increase of temperature because there are more collisions between particles and more energy is produced due to friction.
	33) An adiabatic compression of a perfect gas results in an increase of temperature because there is the same number of particles in smaller volume, then particles are more squashed, there are more collisions, and more heat is produced.
	34) An adiabatic compression of a perfect gas results in an increase of temperature because when the piston is pushed down, the volume decreases; therefore, particles are closer to each other, so more collisions occur between them … and there is an increase in temperature.
	35) If an equal amount of heat is transferred to two systems consiting of the same numbers of particles of a perfect gas at the same temperature but in containers of different volumes, the amount of heat is more diluted in the larger container, so the temperature does not increase as much as in the smaller container.
	36) In molecular terms, the pressure increases in an adiabatic compression of a perfect gas because the volume decreases; therefore the molecules are closer to one another; therefore, there are more collisions and the pressure increases.
	37) When cooling down a liquid, particles become motionless without any order; it is an amorphic solid.
	38) At a very high altitude over the sea level air molecules are far less numerous, therefore pressure at such altitude is considerable less than pressure at the sea level.
(Summers, 1983)	39) Heat is the total energy of the molecules moving around.
	40) Heat is the energy of the random molecular motion.

Table A2. Naïve ideas, printed verbatim, that are listed under the heading 'Structure of Matter' in Ideas Previas and that relate to probability.

Source	Naïve Ideas
(Alsop, 2001)	1) The Radioactivity gets into the person's genes and then this results in genetic mutation.
(Cros et al., 1988)	2) There is not interaction between the particles that make up the nucleus.
	3) Nuclear energy can be released during collisions between molecules.
	4) Nuclear energy can be released during collisions between atoms.
	5) Nuclear energy can be released during collisions between nuclei.
	6) Nuclear energy can be released during collisions between nuclei and particles.
	7) Nuclear energy can be released during collisions between atoms and particles.
(De Posada, 1997)	8) The internal structure of a metal is formed by little bits of metal.
	9) The internal structure of a metal is formed by thin and short lamina.
	10) The internal structure of a metal is formed by little stones.
	11) As the light does not pass through metals, molecules must be very close to one another.
	12) The light passes through transparent materials, then the separation among molecules must be very large.
(Fischler & Lichtfeldt, 1992)	13) There is a distance between the shell and the nucleus so that the electron cannot get to the nucleus.
(Griffiths & Preston, 1992)	14) An atom resembles a sphere with components inside.
(Harrison & Treagust, 1996)	15) Electron shells are as shells enclose and protect atoms, while electron clouds are structures in which electrons aree embedded.
	16) An atom has a dense 'structure' in the middle, like a ball, made up of tiny particles that are neutrons and protons, like a raspberry or blackberry. This inner structure is surrounded by electrons, tiny little particles much smaller than the nuclear components.
	17) Atoms are hard.
	18) The nucleus controls the atom's activities.
	19) Atoms would be soft.
	20) An electron cloud is as a matrix which the electrons are embedded like the water droplets in a cloud.
(Hogan & Fisherkeller, 1996)	21) Atoms and molecules are like a circle with dots; they're made of stuff.
(Taber, 1998)	22) Each electron has the same amount of force between the nucleus and itself, no one electron has more force to it than another.
	23) You need more energy to remove the second electron, the charge on the nucleus is still the same, and there´s one less electron to attract, so the nucleus has more attraction to these [remaining] electron.
	24) When an atom was ionised, the electrons would be attracted more, because the same positive charge pulling on less electrons.
	25) As the ion has an extra electron in its valent shell, the core charge, which remains the same, has to spread its attractive forces equally to each electron thus resulting in less attractive forces on each valent electron and larger atomic radius.
	26) When an electron is removed the effective core charge is sheared out between one less electron therefore increasing the energy needed to remove another electron.
(Lijnse et al., 1990)	27) The half life of a radioactive isotope is the period in which loses its radioactivity. During this period is dangerous.
(Millar & Gill, 1996)	28) Radioactive matter can only travel a certain distance through air. This is because the radioactive matter strikes air molecules and is deflected.

Table A3. Naïve ideas, printed verbatim, that are listed under the 'Astronomy' heading of Ideas Previas and that relate to cosmology and probability.

Source	Naïve idea
(Marques & Thompson, 1997)	1) Before the Earth there was nothing at all.
	2) The Earth originated by collision between planets.
	3) Before the Earth there was something like the Garden of Eden.
	4) Before the Earth there was a vacuum with fire and few small particles.
	5) A great explosion was responsible for the origin of the Earth.
(Sharp, 1996)	6) The Sun smashed into something and it made the planets.
	7) The universe origin is the Big Bang millions of years ago. All planets and things in the universe was one big lump. All the gases made the universe explode.
	8) The planets are made from little bits of rock flying about which joined together.