

Dissecting an interview about astronomical distances using three models

Willem Keppens[✉]

*Department of Physics and Astronomy, KU Leuven, Celestijnenlaan 200C, 3001 Leuven, Belgium;
LESEC, KU Leuven, Belgium;
and LIVO, KU Leuven, Belgium*

Christopher Robin Samuelsson[✉] and Urban Eriksson[✉]

Division for PER, Department of Physics and Astronomy, Uppsala University, 75120 Uppsala, Sweden

Lauren Barth-Cohen[✉]

*Educational Psychology and Physics and Astronomy, University of Utah,
1721 Campus Center Drive, SAEC 3220, Salt Lake City, Utah 84109, USA*

Jan Sermeus[✉]

*Department of Physics and Astronomy, KU Leuven, Celestijnenlaan 200C, 3001 Leuven, Belgium;
LESEC, KU Leuven, Belgium; LIVO, KU Leuven, Belgium;
Faculty of Psychology and Educational Sciences, KU Leuven,
Dekenstraat 2, 3000 Leuven, Belgium;
and Royal Observatory of Belgium, Planetarium, Bouchoutlaan 10,
1020 Brussels, Belgium*



(Received 4 June 2026; accepted 28 August 2026; published 1 October 2026)

There are few studies that compare several theoretical models concerning conceptual understanding. We hypothesize that different models can provide complementary perspectives. In this work, we analyze the video and audio data from an interview with one student (Mark) using three different theoretical models: framework theory (FWT), knowledge-in-pieces (KiP), and social semiotics (SS). The interview discusses distances and sizes in the local universe (Earth, Moon, and Sun) and observable astronomical phenomena: the apparent sizes of the Moon and Sun, solar eclipses, and lunar phases. The results suggest a misconception (from FWT), a localized p-prim (from KiP), and a transductive chain (from SS). However, much more than the individual contributions, the results suggest the complementary nature of the different models. We argue for more work that combines different models both to understand data in a more detailed and nuanced way and to advance theory.

DOI: [10.1103/9vhd-fbfv](https://doi.org/10.1103/9vhd-fbfv)

I. INTRODUCTION

Investigating students' conceptual understanding and the related conceptual change processes is a key area of interest in education research, in particular in science education and its disciplinary subdomains (e.g., physics education research). A quick search on Web of Science shows that there are more than 10,000 papers on the matter, see Fig. 1.

However, Potvin *et al.* [1] found that there are 86 different models¹ about students' conceptual understanding/change. While it might be possible to reduce this number (by assimilating similar models), a large number will remain, which suggests that the field of conceptual change has persistent unresolved research questions. In other words, we, the science education research community, do not really know how conceptual change happens.

¹Others use framework, theory, approach, perspective,... While they may have a specific meaning, they are used interchangeably in the literature. To avoid confusion with framework theory, we decided not to use either of those terms and follow Potvin *et al.* and refer to everything as a model. Not to be confused with mental models (that will always be referred to as such). We will use approach to refer to methodological aspects.

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

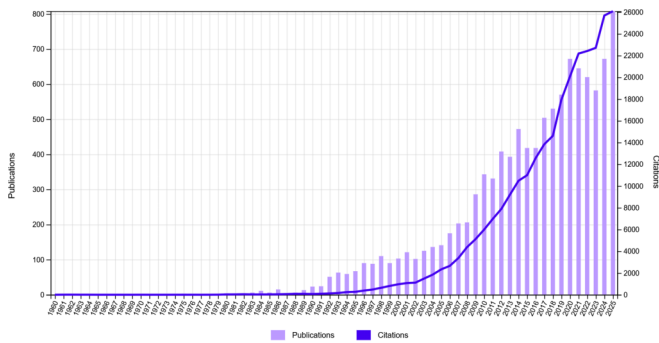


FIG. 1. Web of Science citation report of the search [TS = (“conceptual change”) OR TS = (“conceptual understanding”)] in December 2025.

Indeed, there are fundamental debates that remain unresolved despite more than 30 years of sustained research. Of particular interest is the so-called coherence debate. On one side there are those that suggest that students construct (semi-)coherent mental models, potentially based on fundamental framework theories [2–4]. On the other side, there are those that suggest that students have a complex knowledge system, and when presented with a problem or phenomenon, combine various knowledge pieces and other resources [5]. The conceptual understanding of the student is thought to be dependent on the situation and thus not coherent.

In recent years, the different models in the coherence debate have been converging [6]. diSessa *et al.* [7], a leading author on the knowledge-in-pieces (KiP) side, has suggested that there is at least some coherence.

Debate over coherence vs fragmentation is subtle. No one thinks children are completely unsystematic in their thinking about physics or biology. Furthermore, all existing views of scientific competence, when actually achieved, entail substantial systematicity. The central issue, rather, is specification of the nature and extent of systematicity. (p. 14)

Vosniadou [8], a leading author in the framework theory (FWT) model, has suggested that conceptual understanding in students is at least partially incoherent.

The framework theory approach (Vosniadou, 2013) does not exclude the possibility that knowledge elements such as p-prims might be present in our knowledge system. (p. 2)

Still, despite efforts to converge toward more encompassing models [9], there seems to be a divide regarding the structuring of the mental knowledge structures, which does not seem to get resolved. This is, at least in part, due to methodological issues. Research that is theoretically based in FWT often involves large N , young children and

concepts that are commonly encountered in elementary-level science education, such as the shape of the earth. The KiP side, on the other hand, often opts for individual participants, higher education, and topics associated with introductory-level physics classes (e.g., mechanics).

Another methodological consideration in this debate centers on the relationship between a student’s internal understanding of the physics and how they communicate that understanding. Indeed, it might be impossible, by analyzing a participant’s utterance, to assess their conceptual understanding, as it relies on the student’s mental model, which is fully internal to the participant [10]. Hence, it might be of more interest to look at the way students communicate their meaning. Here, social semiotics (SS) [11] provides another attractive model to use in the analysis of students’ explanations.

Despite a large literature base and foundational papers that have been cited 1000+ times, there exist only a few studies that have directly and empirically compared FWT with KiP, and there are no papers that directly compare SS with classical conceptual understanding models (see background). Yet, such comparisons are essential to move the field forward given the importance of students’ expressed meaning in our data. Either it reveals that one model is superior to another (i.e., better at explaining observed behavior and better at making predictions), or more likely, such a comparison will make apparent where the three models are complementary. As such, there is a need to directly compare theoretical models. In this work, we will do exactly that. We will analyze one interview with a single student, talking about basic astronomical phenomena, from three models: framework theory, knowledge-in-pieces/resources, and social semiotics.

This work is the result of a serendipitously overlapping research stay at the “Physics and Astronomy Education Research” group (Uppsala University, Sweden, chaired by Urban Eriksson) by both Jan Sermeus and Lauren Barth-Cohen. The idea of this work grew during *Fika*, typical Swedish coffee breaks. The choice of these three models is thus in part due to the intellectual respect and curiosities that grew from our *fika* discussions. Still, we think using (these) three models to interpret the same data (not similar or replicated data, which would be subject to at least small contextual differences) could be powerful for building a new understanding of the similarities between and differences across the three theoretical models in a more nuanced and specific way that goes beyond prior methodological differences.

Working together across theoretical models is not new. Kersting *et al.* [12] reported conversations across theoretical divides on embodiment. Their approach reexamined existing data, taken from the literature. Inspired by this, we too were looking for a data-driven approach, ideally using data that came out of an agreed-upon methodology. One option is “adversarial collaboration,” a term first coined by

Kahneman and co-workers [13]. In it, parties that disagree come together to design and run an experiment to test the aspect they disagree on. We too came together from different theoretical stances to look at the same data. This eliminates slight variations in context that can be expected if similar data are used (e.g., data collected by each partner in their own context). However, unlike adversarial collaboration, we did not come together to eliminate but rather to look for complementarity. We did so because we do not think the three models are exclusive to one another. Since adversarial collaboration does not describe what we did, we will refer to our approach as “coopetitive collaboration” (referring to coopetition [14], a noun that refers to cooperation between competitors).

The paper has a traditional structure; it is thus comprised of a background section, where we discuss the direct comparisons of models, a methodology section, results, discussion, and conclusion. However, as will be explained in the methodology, we will report on this comparison through three “mini-papers.” The results of this work are thus three sections that each contain a “normal” paper structure (from an introduction and theoretical background through methodology and results to discussion and conclusion). To avoid confusion, these mini paper subsections are labeled with FWT, KiP, or SS, respectively.

II. BACKGROUND

A. A line of replication studies

In 2004, a decade after the foundational works in FWT [2] and KiP [5] were published, diSessa *et al.* conducted a replication study [15] of the work in 2002 by Ioannides and Vosniadou ([16] from here on denoted as I&V2002). They were not able to replicate the results found by I&V2002, nor were Clark *et al.* [17], who expanded the comparison to other countries. The results from this line of replication studies showed that students were less consistent/coherent than initially reported by I&V2002. This was interpreted in favor of the KiP model by the respective authors, all of whom were working from the KiP model. This interpretation was not shared by Vosniadou and Skopeliti [18]. Rather, they argue, “there is often disagreement on the exact number of students who can be assigned to internally consistent models” (p. 1435) because “the distinction between these two types is not a dichotomous one but rather a continuum with clear cases at the end and some difficult to distinguish cases in the middle” (p. 1435).

However, as diSessa *et al.* argued [15], perhaps the difference is not a matter of interpretation but more so related to the fact that they did not conduct a true replication. Indeed, they refer to it as a “quasireplication.” It is perhaps impossible to truly replicate a study in educational science since there can be a participation effect where the participants learn something by participating (in addition to aging, additional learning, and life

experiences, among other factors). This is, however, not the problem with the replication they refer to. Instead, the studies were conducted in different countries, each with its own curriculum, culture, and language. Indeed, it has been suggested that these could have a large impact on students’ understanding of concepts [19]. This makes it difficult to interpret the replication results. As Clark *et al.* remarked [17], it is unfortunate that the international replication did not include Greece (the study where I&V2002 was conducted).

B. Comparative studies

If not through replication, has there been any effort in the literature to directly compare the different models? To the best of our knowledge, there is very little.

There has been work by diSessa [20] comparing KiP with Theory Theory (which encapsulates FWT, among others) and the Ontological View (referring to the work of Chi *et al.*, e.g. [21]). Perhaps not surprisingly, he found that KiP worked well to provide “a more detailed and more adequate account of the learning that took place, whereas that learning contradicts core commitments of the Theory Theory and of the Ontological View” (p. 1).

In 2000, Elby [22] suggested that the KiP (referred to as “fine-grained constructivism”) and FWT (referred to as “misconceptions constructivism”) approaches could be distinguished experimentally given a careful research design. The central argument is that

FWT predicts stable mental constructions that can switch during conceptual change but with nothing happening in between, while in contrast the context-dependent cueing of [KiP] allows stories to be told about the fine structure of these transitions, stories that predict particular patterns of nonrandomness in the fluctuations. So, [KiP] generates predictions about students’ behavior in cases where misconceptions constructivism predicts nothing other than random fluctuations. (p. 491)

Elby presents three pilot studies where students are confronted with a visual that could trigger a “What-you-see-is-what-you-get” resource. Elby finds that more students than randomly expected make the same mistake, which he interprets as evidence in favor of KiP.

In contrast to Elby, the work of Blown and Bryce [23] “supported conceptual coherence and rejected the counter-argument of knowledge-in-pieces (at an alpha level of .05)” (p. 31). This was the result of a 10-year longitudinal study with 345 young participants from both New-Zealand and China. They investigated the participants’ understanding of the astronomical concepts of *seasons* and *eclipses*.

A final paper that directly compares both models is the work of Scherr [24]. She compared the same data using

both FWT and KiP. Her analysis showed that the data could be interpreted by both models. In the summary, she writes [24]

The available data about student reasoning in special relativity favors both misconceptions and pieces models of student reasoning—sometimes one, and sometimes the other. Regarding the fact that simultaneity is relative, student ideas are more consistent with a misconceptions model (or with the misconceptions-like part of a pieces model). Students learning about the nature of a reference frame, on the other hand, behave more as predicted at the other end of the spectrum of possibilities in the pieces model. (p. 10)

As such, she concludes that the two models are complementary and that researchers and teachers should be open to both. She was, however, not explicit about where the models are complementary.

C. Studies with complementary models

The call of Scherr to use FWT and KiP as complementary models is nontrivial, though. There are, to the best of our knowledge, few authors who tried to interpret data from multiple theoretical perspectives.

In 2006, Clark [25] reported on a longitudinal study following 50 grade 8–12 students as they learned about thermal equilibrium. The study did not directly test the FWT or KiP model; rather, it reports on the learning of the 50 participating students and in detail about the learning of four case-study students. His analysis focused on several grain sizes: topic, explanation, and elements within explanations. Clark discusses that the data could be described well from a KiP perspective (or, as he refers to it, “elemental perspective”). On the other hand, even if a theory-like character is assumed for the students’ initial ideas, as they learned “it seems quite possible that many of the students will remain in a state of flux rather than achieve a parsimonious, coherent, theory-like understanding of thermodynamics throughout adulthood” (p. 547). Clark suggests that “theory-like perspectives needs to consider the extended stretches of time between these theory-like periods” (p. 548).

Rodriguez and Towns [26] analyzed biochemistry students’ graphical reasoning from both a “misconceptions constructivist” and “fine-grained constructivist” approach, in line with Elby [22] referring to FWT and KiP, respectively. They argue that by taking on a different theoretical stance, researchers will set up different research questions, run different experiments, and come to different conclusions. With an FWT lens “the logical approach for interpreting data is to evaluate whether these concepts align with scientifically accepted ideas and focus on how to address a lack of alignment” (p. 13). On the other hand,

KiP naturally leads to studying how “how context influences emergence of ideas as well as the organization of ideas within a manifold structure” (p. 13). While their results show interesting insights that came from both ways of interpreting the data, the authors seem to prefer KiP.

Finally, we found some authors who actively switched from one model to another. For example, Farlow *et al.* and Vega *et al.* initially started their “research project [...] guided by a framework consistent with ‘identifying student difficulties’” [28] (p. 194). However, they switched to a resources-based model because “students’ responses were not consistent with stable, strongly held ideas; rather, student responses seemed to be more fluid and context dependent. In addition, incorrect answers involved elements of correct thinking, and a resource model allowed us to make use of the productive resources that students access and build upon them, as well as to identify unproductively applied resources.” Such switching of models shows the flexibility of the authors. However, rather than complementing one approach with another, by adhering to one model strictly and neglecting the other, any possible complementarity is lost.

D. The body as a bridge between mind and environment

While the ideas of conceptual change were developed with the learner’s cognitive processes in mind, social semiotics has focused on representations, the activities, and tools for communication in the environment in which the learning takes place. The mind, and in particular the body, serves as the primary bridge between these two areas. However, only a few attempts have been made in integrating [29,30] or contrasting [2,31] the focus on the cognitive processes and the social activities despite the assumed influence of one another across different “grain sizes” [29].

In more recent years, this “bridging” effort has gained momentum [6], focusing on the central role of the body in meaning-making. Kersting *et al.* [12] illustrate this by reanalyzing a single sequence of data through cognitive, interactionist (similar to social semiotics), and phenomenological lenses. They argue that the body serves as a bridge by functioning simultaneously as a “role player,” providing the “image schemas” (e.g., balance, containment, and rotation) that ground internal mental models, and a “sensor” that interacts with external semiotic resources (e.g., gestures and material tools) to coordinate shared disciplinary meaning. At the same time, from a social or interactionist perspective, the body acts as a powerful semiotic resource, where gestures and whole-body performances are coordinated alongside language and diagrams to construct shared disciplinary meaning.

This multilayered perspective suggests that conceptual change is not merely an internal mental reconstruction. This allows for a discussion of the role that semiotic

resources, like lab equipment, play in the cognitive process of the learner [32]. Furthermore, Halldén *et al.* [33] have proposed a methodology (“intentional analysis”) that interprets student utterances as “situated actions,” allowing researchers to infer stable conceptual structures while accounting for the immediate context.

Building on these theoretical and methodological bridges, the question of whether a full reconciliation between these traditions is achievable remains a central provocation for the field. As Mason [34] argues, despite the “warming trend” in cognitive research [35,36] that has begun to incorporate motivational and situational variables, there is still a pressing need to move beyond seeing sociocultural perspectives as mere “replacements” for cognitive views. Instead of falling into the trap of dichotomizing these perspectives, Mason suggests that a truly comprehensive account of learning requires a finer-grained understanding that employs both cognitive-information processing and sociocultural lenses simultaneously to capture the multiple facets of the conceptual change process.

A promising methodological avenue for this integration is the “networking of theories” approach, as exemplified by Svensson and Campos [37]. By coordinating social semiotics with more cognitively oriented semiotic frameworks to analyze the same dataset, they demonstrate that a wider interpretation of student learning is possible. In their model, social semiotics provides a descriptive account of how students discern and manipulate disciplinary resources, while the cognitive lens captures the internal construction of knowledge and the dissociation of concepts from their specific representations.

While recent work on the bridge of embodied cognition and other methodologies offers promising ways of integrating or comparing cognitive and sociocultural models, the feasibility of a unified model that accounts for both the internal processes of the mind and the sociocultural genesis of knowledge is still an open area of inquiry. Consequently, further research is needed to develop coherent research programs that do not just treat these as parallel processes but as an integrated “knowledge in context” [33]. Such research, and specifically, the comparative analysis of the same data sequences, is essential to determine how individual cognitive structures and collective discursive practices co-evolve, ensuring that the full complexity and multimodality of science learning are understood.

III. AIM

From the background, it is clear that we found a few studies that compare several models and that those that do come to different conclusions. One idea might be that the models are complementary, yet we also found that there are few studies that seek to exploit this potential complementary nature of the different models. Moreover, those

that do tend to fall back into one model. It is clear that more work is needed to cross the theoretical divides and to directly compare several models. The aim of this paper is thus to compare, using a single dataset, three models: social semiotics (SS), knowledge-in-pieces/resources (KiP&R) and framework theory (FWT). All models are interested in students’ ideas and how they express them. We think this comparison is theoretically interesting, showing where the different models converge, where they are complementary, and where they are contradictory, while also pointing to new research directions that move beyond historical differences. Moreover, we think a direct comparison might be of interest to early-career researchers in science education.

IV. METHODOLOGY

A. Interview procedure

All three theoretical models value interview data for the flexibility they offer in allowing the interviewer to follow up on the interviewees’ key statements and explore their reasoning further [38]. The FWT approach tends to implement an interview approach closer to a formal interview style with an emphasis on predetermined questions, while the KiP approach tends more toward a mixture of predetermined questions with an informal approach of some unstructured conversations [39] and the SS approach necessitates space for multiple modes. Despite these differences, we will work with interview data in this work.

The interview procedure and the preceding survey, described below, follow that of the work reported in Refs. [40] and [41]. That work is conducted, in part, by the first and last author and was developed from the framework theory model. However, all authors agreed on the interview procedure explained, and, importantly, all authors agreed that this procedure had the potential to lead to data that would be of interest to their model. While an interview designed specifically within any model might be slightly different than within another model, in general the questions asked, the flow of the interview, and the props used were agreed upon to work for all three models given their approaches to data collection.

These data consist of (i) survey results and (ii) interview data that were video recorded. The survey asked participants to estimate sizes and distances in the local universe (Earth, Moon, Sun, next Star), as well as answer multiple-choice questions that probed the students’ conceptual understanding of seasons and Moon phases. It is an extension of the survey as described in Ref. [42]. This survey was administered, in English, to one class in a Swedish secondary school with students of ages 16–18 in general track education (though the school had a more technical and natural sciences focus).

From this class, three students were selected for an interview 2 days later. Students were selected if they

answered in the survey that the Earth, Moon, and Sun were relatively large and/or close to one another, i.e., if they imagined the local universe as “dense”. This selection means it would be more likely for students in the interviews to express erroneous understanding (uncovering this was the purpose of Refs. [40] and [41]). It also made it possible to present a full-scale model of the students’ Sun-Earth-Moon system on a table during the interview that followed.

During the interview, the students were presented with their personalized scale model and asked if they would change anything. This scale model, a two-dimensional model where the Earth, Moon, and Sun were drawn on paper/post-it notes, was on the table during the whole interview, and students were encouraged to use it in their explanations. This scale model needed to be flexible and practical; that is, students needed to be able to replace, resize, and/or redraw the celestial objects in the model at will. Consequently, 3D models (for instance, using balls or balloons) were not suitable for this research, although we acknowledge that their use would have been better at promoting students’ 3D spatial thinking.

The interview covered three topics. First, the students were asked to estimate the size of the Moon and Sun as seen from Earth. They were offered the choice of a series of discs mounted on sticks that they were to hold at arm’s length. They were also asked if this size might change over time. Second, the students were asked to explain solar eclipses. Here students were, after an initial answer, confronted with two photographs of a full and an annular eclipse and asked to explain the possible difference. Third, the students were asked to explain the lunar phases. The students were, after their initial answer, confronted with a photo of a three-quarter (gibbous) Moon and were asked to place the Moon of their scale model, such that it would lead to this Moon phase as it would be observed from the Earth in their scale model. All three interviews, each approximately 25 min long, were video and audio recorded and transcribed verbatim.

Finally, from these interviews, one was selected for this shared project: Mark’s interview. From this interview, one section was selected to focus on. The verbatim transcript of this section is given in the shared results. This selection was done based on discussions between the author teams after all authors watched all three interviews. The selected interview was deemed the most interesting and most likely to lead to valuable insights.

Some elaboration is warranted on how the interview protocol was influenced by the framework theory from which it originated. As a critical response to Vosniadou’s findings, Schoultz *et al.* [31] wrote that “one has to be aware of the fact that we tend to organize data collection in such a manner that we find the entities that we are looking for.” We agree with this statement to some extent. As FWT focuses heavily on the consistency in students’ answers

and the coherence of their mental models, the interview questions indeed were formulated with this focus in mind. For instance, after students had chosen a Moon or Sun disc to express their estimate of its apparent size, they were asked if they thought that the disc was compatible with their scale model on the table. In other words, would someone on the scale Earth observe an equally large Moon or Sun in their sky, based on the scale model’s proportions? This question probes the students’ consistency between different answers (the chosen disc and the constructed scale model) and their self-evaluation of this consistency (do they judge those two modes to be compatible?).

However, this series of questions can also be of interest and can be interpreted in a valid way by other models. For instance, one could argue that it probes whether the introduction of new (semiotic) resources (such as the scale model) alters students’ reasoning. We therefore pose that the adopted model indeed can influence interview protocols. However, logically, this does not mean that if one adopts a different model, it necessarily changes the interview protocol (or indeed the answers of the students).

We chose to use many different tools and artifacts in the interview protocol, such as the scale model, the discs on sticks, and pen and paper for students to make drawings. This is not in line with the classical approach of Vosniadou, who used no tools except for drawing paper when interviewing students [2,3]. Research has shown that the inclusion of tools or artifacts can engage students in much more sophisticated modes of reasoning, thereby drastically altering the results of an interview study [31]. As such sophisticated reasoning patterns were of particular interest to this study, we wanted to provide all potentially useful tools for students to express them. Moreover, we deemed artifacts such as the discs on sticks to be essential for students to express their estimates on a topic as specific and abstract as the apparent Moon and Sun sizes. As such, even though the interview protocol was rooted in framework theory, not every aspect of it was tuned accordingly.

B. Mini-papers

We analyze one common piece of data—an interview with a student—but the data are analyzed in three different ways, corresponding to the three models. To allow a direct comparison, this analysis will be reported as three “mini”-papers where the author teams representing each model detail how they process the same data. These papers are maximally synchronized, each containing the same typical sections one would find in an academic paper. We note that, for brevity, the theoretical background will not include references to other models, the methodology will only discuss the analysis (data collection is in Sec. IV A), and the results will point the reader to specific results (all results are in Sec. V B).

V. SHARED RESULTS

The outcome of Mark's scale-assessment survey, the artifacts, a video and audiorecording of the interview, and the transcript² of the interview served as the data that were analyzed by the different author teams. Here we present these data as condensed as possible but as extensive as necessary. We share the full transcript of one section of the interview and summarize Mark's answers to other questions.

A. Scale model

The physical scale model of the Sun-Earth-Moon system represents a central element of the interview context. Constructed from Mark's answers to the preceding online survey, it plays a key role in the analyses of all three models. Therefore, we first shortly describe the dimensions of the scale model on the interview table.

Mark indicated that the Moon should be relatively close to the Earth, with only one quarter of the Earth's diameter in between the two bodies (the actual distance between Earth and Moon is roughly 30 Earth diameters). A consequence of such a close-by Moon is that the subtended angle by the Moon (i.e., the angle as an observer would see it from the scale model Earth) is very large, approximately 30°.

With regard to the Sun, Mark estimates the diameter of the Sun as roughly 800 times larger than that of the Moon and places the Sun at "only" 200 times the distance between Earth and Moon. This results in the Sun's subtended angle being even larger compared to that of the Moon.

Mark was presented with a physical version of his scale model. The Earth and Moon were drawn onto post-it notes, respectively, as a 8 and 1 mm circle. The Sun was drawn on an A2 paper as a 58 cm circle (according to Mark's survey answers, it should actually have been 80 cm. But this was not feasible given the available space). This physical model was on the table as he entered the room and was present throughout the interview.

B. Verbatim transcript

In this section, we state the verbatim transcript of one particular fragment of Mark's interview. It is a fragment in which Mark thinks, introduces new ideas, and uses his hands and other resources to explain his ideas. This fragment is preceded by the first 5 min of the interview, which consists of a welcoming message and the initial questions where Mark has to estimate the apparent size of the Moon and Sun in the sky. It is a section with less thinking and is thus of less interest (though, of course, it is necessary to set up the rest of the interview). In the rest of

the interview, the 13 min that follow this transcript, Mark answers various other questions but continues to build on the ideas that were first discussed and uncovered in the 5-min fragment transcribed below. As such, this fragment stood out to all three teams as interesting.

In the excerpt below, **I** indicates the interviewer, and **M** stands for Mark. We also enumerate the line number of each sentence, as they are often referred to in the following sections. Filler words are included in the transcript, as they might indicate thinking or doubt (and thus are indicative of an (un)stable mental model). Square brackets ([]) are used to describe gestures made by Mark or by the interviewer, which is necessary to allow an analysis from social semiotics.

- 47 I: Do you think that the size of this Moon to this person [points at the scale model] is the same as the size that you just picked [points at the chosen Moon disc]? Or is there a difference between the two?
- 48 M: Euhm, well...
- 49 I: Do you understand the question, or..?
- 50 M: I, like, do we perceive it differently depending on where on earth we're standing or..?
- 51 I: No, that's not really the question.
- 52 M: It's just like if you're standing here, is this how it's going to look?
- 53 I: Yeah, exactly, is this [points at the chosen Moon disc] how it's going to look?
- 54 M: Yeah, uh, well I mean here in Sweden, we're kind of close to, you know, the top of the Earth so I, yeah, I'd say so, yeah.
- 55 I: Okay. And do other people on Earth see other sizes then?
- 56 M: Yeah...oh wait, actually, I think if you stand on the top [points at the model Earth], you know it's...It's a bit more, you know, further away compared to other parts. So I think the sizes would, would kind of like change.
- 57 I: Okay.
- 58 M: But I think... yeah, no, I think I think the size would change and, yeah.
- 59 I: So if I'm on the equator like, right here [points at what would be the equator on the model Earth], it would be different?
- 60 M: Yeah.
- 61 I: Okay. How much different do you think? Like would you pick another of these Moons [points at the other Moon discs] if you were on the equator?
- 62 M: I'd say like maybe just like...small incremental change maybe, maybe like this [holds a slightly larger Moon disc next to the originally chosen one]?
- 63 I: Okay. For the Swedish observer, compared to the equatorial observer?

²The transcription already implies some interpretation, this is discussed in Sec. VB.

- 64 M: Euh, yeah. So, so the bigger one would probably be the equatorial one and the smaller one would be closer to Sweden.
- 65 I: Okay. And if you would stay in Sweden, but you would observe the Moon at different times. Would there then be a variation?
- 66 M: Well, I mean it would technically go further away or closer to, you know, where I'm standing, you know, when in its orbit. So yeah, I think, I think it might you know also like change a little bit, but not, not as noticeably as, to like watch it at the equator compared to... yeah.
- 67 I: Okay. So there is a small difference you say, but is the difference in this case, while staying in Sweden, also large enough to pick like a range of these disks? Or would you say no, this disk is always the best approximation?
- 68 M: Euh, around this disk, I would say.
- 69 I: Alright. And so, to go a bit further on the why, so you're saying that it has something to do with the orbit?
- 70 M: Euh, yeah, oh wait no, I mean it depends on like if the Moon orbit is like a complete circle [moves his hand in a circle shape above the table] or if, you know, it gets shorter in some spans or longer in some spans [moves his two hands repetitively closer and further from each other].
- 71 I: And what do you think is it?
- 72 M: Euh, I think it isn't a complete circle. Euh, but I'm not sure obviously. So, you know, sometimes when the Moon is around the span between me and the Moon would then be shorter than other parts of, you know, between [holds his left hand in front of him and moves his right hand back and forth between his body and his left hand, indicating the distance in between. Then rotates his body and his left hand to the left side, and makes the same gesture with his right hand again].
- 73 I: Can you make a drawing of that [points to the blank paper on the table]?
- 74 M: Sure. It's like maybe a...so let's say this is Earth, E for Earth [draws a circle with a letter E inside of it, see Fig. 4].
- 75 I: Yeah.
- 76 M: And we have the Moon and it's kind of...big, bit bulgyer here maybe [draws the Moon orbit].
- 77 I: Yeah.
- 78 M: Yeah. And then you have the Moon [draws a Moon], and let's say, here I am [draws a person on the Earth].
- 79 I: Yeah.
- 80 M: So here we have one span [indicates the distance between the person on Earth and the Moon], let's say it's, it's a little bit closer here [draws a second Moon at another position in the orbit, using a dotted line], like if the Moon was...here.
- 81 I: Which one's closer?
- 82 M: Euh, the dotted one. [points at the second, dotted Moon in the drawing]
- 83 I: Okay. And that's due to the orbit of the Moon?
- 84 M: Yeah.
- 85 I: Okay. So we have the effect of the orbit of the Moon and we have the effect of the position of the observer you say, right?
- 86 M: Yeah.
- 87 I: Okay, all right. And do you actively remember instances where you looked at the Moon and you thought, wow, now it's surprisingly large, or maybe surprisingly small?
- 88 M: Euh, not, not really. I honestly, I don't look at the Moon that much, I mean it's mostly...the thing I noticed about the Moon is mostly just how one part of it is shaded and one part of it's, you know, if it's a half Moon, if it's a full Moon or, you know.
- 89 I: Yeah.
- 90 M: Yeah, the faces of the Moon.
- 91 I: Yeah.
- 92 M: But I never, I never really think about the size.
- 93 I: Okay, no problem, very well. Let's leave these Moons for what they are for a moment and instead turn to these Suns. So also I've brought this range of Suns with varying sizes [points at the Sun discs].
- 94 M: All right.
- 95 I: Imagine you're going outside right now and you have these protective sunglasses and you look at the Sun.
- 96 M: Yeah.
- 97 I: Can you then again, try to compare what you would see then with what you see when you hold these discs, again at arm's length [illustrates by holding one of the Sun discs at arm's length]?
- 98 M: All right, yep.
- 99 I: And then try to pick one again.
- 100 M: All right, let's start with this one [picks one Sun disc and tries it]. Yeah, 'cause there's the Moon so... it's a lot further away, but still a lot larger [puts the Sun disc down on the table], so it would still be in the same consention, I mean, it wouldn't be one of these two [points at the two largest Sun discs], obviously. And I'm also skeptical about these [points at the third and fourth largest discs], so...it's got...this one [tries another, smaller disc]. There's no middle one between these two [points at the firstly and secondly tried discs], no?
- 101 I: Hmm, let's see [searches quickly between the other discs]. I don't think so, no.

- 102 M: Then let's say, let's say about this one [picks up the larger of the two discs and hands it over to the interviewer].
- 103 I: Okay. [looks which disc it is by checking the indication on the backside, and puts it back down on the table]
- 104 M: Yep. This one's closer to...

C. Other observations and artifacts from Mark's interview

In this section, we summarize the other observations of Mark's interview. We only provide a brief and objective description of Mark's answers here, as any more thorough analysis might be subjected to the writer's interpretation. Enumerated according to the topics discussed in the interview, we found that:

- **Apparent sizes:** Mark indicated that he believed the apparent size of the Sun (as seen from Earth) to be larger than that of the Moon. As such, the Sun disc he chooses is larger than the Moon disc. This is in accordance with the scale model on the interview table, in which the Sun also subtends a larger angle than the Moon. However, the subtended angles of both bodies in the scale model are roughly 10 times larger than those subtended by the respective discs (when held at arm's length). The Moon disc was chosen before the start of the above transcript, while the Sun disc was chosen during this fragment (line 102).
- **Variation in apparent Moon size:** Mark assigns two causes to a variation in apparent Moon size over time. The first is the position of the observer. An observer near the poles would be physically farther from the Moon than an equatorial observer and would therefore see a smaller Moon. The second is the non-circular orbit of the Moon around the Earth, which Mark judges to be a smaller effect. This discussion is included in the above transcript excerpt and is made explicit in line 66.
- **Variation in apparent Sun size:** The Sun varies in apparent size over time due to the noncircular orbit of the Earth, Mark argues. For the Sun, the effect of the observer's geographical position does not hold because of the larger distance to the Sun compared to the Moon. This topic was discussed after the above transcript excerpt and was illustrated by Mark in Fig. 2.
- **Solar eclipses:** Mark knows that during a solar eclipse, the Moon is in between the Earth and the Sun. He thinks that both the apparent sizes of the Sun and the Moon change between a total and an annular eclipse, but that the apparent Moon size has more

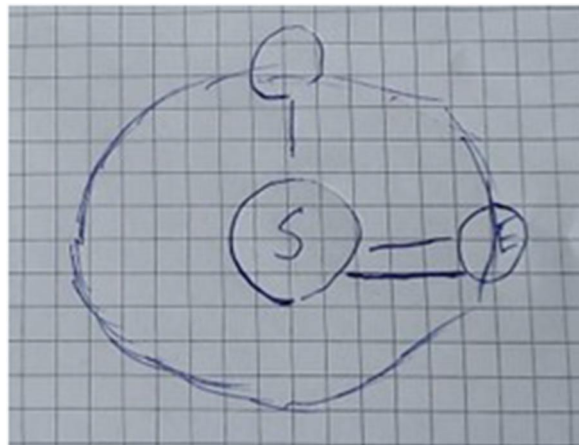


FIG. 2. Mark's drawing of the Sun-Earth system. Mark argues that the position of the Earth in its orbit matters, i.e., that the Sun will appear larger on the right (Earth E) than on the top (unmarked Earth) in the drawing. Mark also argues that the Sun's size does not change depending on where you stand on Earth (indicated by the second line connecting the Earth E and the Sun S).

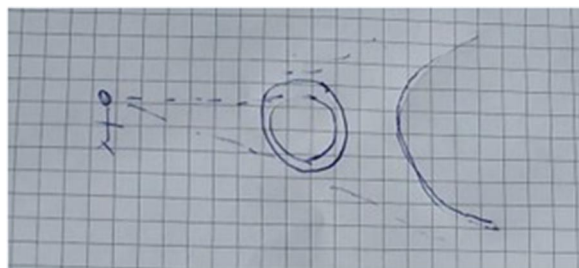


FIG. 3. Mark's drawing of a solar eclipse. The Moon, in the middle, varies more in size than the Sun, on the right.

impact. He argues that since the apparent Moon size relates to the observer's geographical position, a polar observer can see an annular eclipse while simultaneously an equatorial observer sees a total eclipse. This topic was discussed after the above transcript excerpt. Mark made the drawing in Fig. 3.

- **Moon phases:** Mark explains correctly how the Moon reflects sunlight and how the observed phase is determined by the relative positions of the Earth, Moon, and Sun. He accurately links the different phases to their corresponding configuration, including the gibbous Moon shown in a photograph. Mark makes no mention of the Earth's shadow falling on the Moon, which is a common mistake made by students to explain Moon phases. This topic was discussed after the above transcript excerpt. The only drawing Mark made of the Earth-Moon system is shown in Fig. 4.

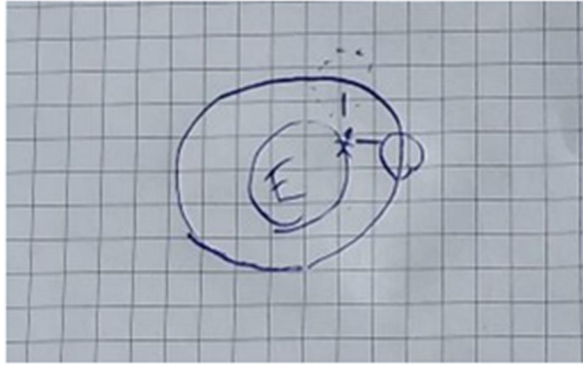


FIG. 4. Mark's drawing of the Earth-Moon system.

VI. MINI-PAPER 1: FRAMEWORK THEORY (FWT)

Authors: Willem Keppens, Jan Sermeus

A. Intro—FWT

By the time students enter the classroom, they have already formed an idea of how the world around them works [18]. This preinstructional understanding is referred to in different ways, from naïve knowledge to preconceptions [43]. Attempts to reconcile this naïve knowledge with targeted instruction in the classroom can lead to the formation of synthetic models [18]. These generally contain elements of the scientific understanding but still explain the concept at hand erroneously. In astronomy education, many such erroneous explanations have been documented. This might be related to an insufficient understanding of the involved spatial scales [44,45]. For instance, students tend to underestimate the distances to the Moon, Sun, and stars [46–48] and they explain the seasons in terms of a varying distance to the Sun [49–52]. However, little is known about how students' understanding of astronomical scales impacts their explanations for astronomical phenomena besides seasons. In this work, we investigate the interview with Mark as a case study. We attempt to build a mental model as Mark might see it in his mind's eye and look into the importance of scale assumptions on Mark's reasoning about several astronomical phenomena. We analyze and interpret the results from the perspective of framework theory.

B. Theoretical background: The framework theory approach to conceptual understanding and conceptual change

The framework theory starts from the idea that before any formal instruction, children have already formed a naïve understanding of physics [18]. This understanding does not consist of independent observations but rather forms a loose but relatively coherent conceptual system—a framework theory. Children have developed this theory

under the influence of lay culture, together with perceptual and sensory experiences.

Unlike scientific theories, a framework theory is not well formed, nor is it explicit or socially shared [53]. It is called a theory because it is a principle-based system, governed by ontological categorization and causal relations, which generates explanations and predictions. For instance, when an entity like the Earth is classified in the ontological category of physical objects (instead of astronomical objects), it will automatically inherit all properties tied to that category, such as an objective up-down organization, a need for support, and solidity [53].

Within the framework theory, the process of learning requires many conceptual changes, including changes in ontological categorization, representation, and epistemological beliefs [53]. Therefore, learning is not regarded as a sudden paradigm shift where erroneous theories are replaced by scientific ones. Rather, it is a slow and gradual process, giving rise to synthetic models and fragmentation [18]. These originate from attempts to reconcile new, often counterintuitive, scientific information with already held and incompatible conceptions of the framework theory. The product of this incorporation is considered a synthetic model when it combines, in an erroneous way, elements from both the scientific and the naïve understanding. While they are scientifically incorrect, synthetic models nonetheless provide some degree of internal consistency and explanatory value. Fragmented conceptions, on the other hand, are those with an *ad hoc* application and without internal consistency [18]. This distinction does not mean that synthetic conceptions are always static and unitary; they are in fact malleable and susceptible to changing contexts. They are not characterized by flawless coherence but rather by a search and desire for coherence and explanatory power by the individual holding them [53].

To conclude, we elaborate on what is meant exactly by the notion of a mental model. In the framework theory, a mental model is understood as a mental representation with an analog character, that is, one that preserves the perceived structure of the thing it is supposed to represent [3]. For instance, a mental model of the Sun and the Earth includes their relative sizes (preserving spatial structure), and a correct mental model of the day-night cycle entails a spinning Earth (mirroring causal structure). Mental models are dynamic and can be manipulated to provide solutions for certain problems at hand or to formulate potential explanations for physical phenomena. As these solutions or explanations may very well be at odds with scientific viewpoints, they could in those cases be categorized as synthetic models. Vosniadou and colleagues have described several alternative (i.e., nonscientific) mental models, for instance, on the shape of the Earth [2] and the concept of number [4].

C. Aim—FWT

We aim to investigate how Mark explains the astronomical phenomena discussed during the interview. We will focus mainly on the reasoning Mark follows when talking about variation in the apparent size of the Moon, although we will also shortly touch upon other covered topics. The main research questions are as follows: Does Mark have a mental model of the Sun-Earth-Moon system, and if so, how (if at all) does Mark use his mental model to explain astronomical phenomena? And which core features of the framework theory—such as a search for internal consistency, overall coherence, explanatory power, and the occurrence of synthetic models—can we identify in this process? Answering these questions could be insightful for mapping Mark's internal knowledge structure, as well as specific mental models. In turn, such insights could lead to a better understanding of the process of conceptual change.

As such, we aim to narrow down the gap in the literature between students' comprehension of spatial scales and their astronomical conceptual knowledge. Such work is important to uncover a potential link between research findings on both topics, which we believe to exist based on our previous work [40,41].

D. Method of analysis—FWT

The interview with Mark was transcribed verbatim, including descriptions of the gestures and drawings made by Mark, as well as his interactions with the scale model. The coding of the data was done on the level of the questions, including potential follow-up questions and subsequent discussions. Thus, the grain size units to be coded were stand-alone thoughts or responses, and no coding of smaller units (e.g., individual sentences or statements) was performed. For instance, toward the end of the interview, Mark is asked to explain Moon phases. He spends slightly over 2 min formulating his reasoning, which he illustrates by using the scale model and moving the Moon post-it around the Earth post-it. All the statements, gestures, and scale model interactions during those 2 min were collectively assigned one code (in this case, the code “AB.1.: correct and complete explanation”). The coding schemes we used were identical to those constructed in our other research projects [40,41]. In these references, the complete schemes are displayed. The codes in those schemes originate either from the existing literature in astronomy education research or from empirical findings in our own data. The analysis method was thus one of theory-driven thematic analysis [54], in accordance with the characteristic methodological approach of Vosniadou's framework theory [3].

As in the previous example of Moon phases, the answers expressed by Mark were interpreted together with the scale model on the table. Special attention was paid to the (in)consistency between Mark's statements and his

(interactions with his) scale model. Consistency between these two means of communication is interpreted as an indication for a coherent underlying mental model, while inconsistencies suggest an incoherent or even absent mental model. On the same account, consistency between several (verbal) answers of Mark also cued the nature of an underlying mental model.

E. Results—FWT

1. Apparent Moon size variation

When discussing potential causes for variation in the apparent size of the Moon, Mark argues that it depends on the position of the observer. He states that “here in Sweden, we're kind of close to, you know, the top of the Earth” (line 54) and that “if you stand on the top [points at the scale model Earth], you know it's... it's a bit more, you know, further away compared to other parts. So I think the sizes would, would kind of like change” (line 56). Mark thus reasons that there would be a larger distance between the observer and the Moon when standing close to the “top” of the Earth, for instance in Sweden, than when standing on “other parts” of the Earth, presumably closer to the equator. Therefore, an equatorial observer would see a larger Moon than a Swedish observer. By making this argumentation, Mark implicitly assumes the orbital plane of the Moon to coincide with the equatorial plane of the Earth. An illustration of this argument is presented in Fig. 5. After Mark claimed this difference in apparent size to be noticeable, he pointed out a second, larger Moon disc, in addition to the one he chose earlier in the interview. Evaluating the two different discs, he concludes that “So, so the bigger one would probably be the equatorial one and the smaller one would be closer to Sweden” (line 64), in line with the mental model in Fig. 5.

Mark suggests that there is also a second cause for variation in the apparent size of the Moon, which is the noncircular orbit of the Moon around the Earth (line 72). This causes the distance between the Earth and the Moon to change over time, which in turn leads to a variation in apparent Moon size (line 66). However, Mark immediately adds to this that this effect would be “not, not as noticeably

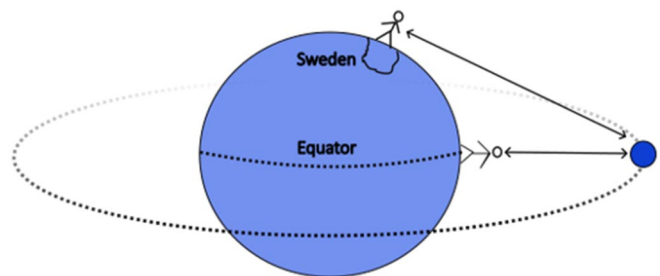


FIG. 5. Mark's mental model regarding the positional dependence on the distance between an observer on Earth and the Moon as interpreted by the FWT authors.

as, to like watch it at the equator compared to... yeah” (line 66), again referring to the mental model in Fig. 5. According to Mark, the ellipticity of the Moon orbit—or in his words, its “bulgyness” (line 76)—has a smaller impact on the variation in apparent Moon size than the geographic location of the observer, in line with Fig. 5.

2. Apparent Sun size variation

According to Mark, the apparent size of the Sun can change as well. However, the variation is now attributed only to the ellipticity of the Earth’s orbit or rather its “ovalish” shape (line 124). He argues that “the length between the Sun and the Earth here [illustrates with a line in the drawing of Fig. 2] would be different than here [with the Earth in another position in the orbit]” (line 126), and further that “in that regard, looking at the Sun from the different distances will show a different sized Sun” (line 128). However, when asked whether the location of the observer plays a role as well, as it did for the apparent Moon size, Mark answers that he does not think so. He says: “I think that has a lot to do with just the.. The length [points at a line between the Sun and Earth in the sketch of Fig. 2] or like the distance between the Sun and the Earth. That it doesn’t change, comparatively, between the placements on the Earth, like the distance between the Earth and Sun here, and the Earth at the equator, and let’s say, well, the.. the south pole...matters a lot less when it comes to the Sun compared to looking at the Moon, yeah” (line 136). While indirect, it reaffirms the mental model in Fig. 5.

3. Comparing apparent Moon and Sun sizes

When comparing the apparent sizes of the Moon and the Sun, Mark picks up his previously chosen discs and concludes that the Sun must look larger in the sky than the Moon. However, when asked whether this was in agreement with what he remembers from experience, he spontaneously refers to solar eclipse events and reasons that in those instances, the Sun and the Moon must be of the same size. He thus argues that occasionally they can be of equal size, when “the Sun is at its smallest and the Moon at its biggest [...] Even though the Sun is usually bigger than the Moon.” (line 146).

4. Solar eclipses

Mark correctly explains the phenomenon of a solar eclipse in terms of the Moon blocking the Sun (as seen from Earth), and he makes a schematic drawing to illustrate his explanation (Fig. 3). However, the described mechanism is far from realistic in the scale model on the table, as there the Sun is much too large to be covered by the tiny Moon. Mark does not notice this inconvenience and is not explicitly pointed to it by the interviewer. When Mark is asked to interpret the difference between a total and an annular solar eclipse, he dedicates the cause to

different geographic locations. He explains: “Well, since the Moon size would be in correlation to where you are on Earth, while looking at the Moon, the, the eclipse would look differently depending on where you would be on the Earth” (line 172). The idea that differently positioned observers see a differently sized Moon (Fig. 5) thus clearly resurfaces in this part of the interview.

F. Analysis—FWT

We first discuss the reasoning of Fig. 5 in more detail. It is interesting to note that despite its incorrectness (or rather, its irrelevance), this reasoning makes a lot of sense when the Moon is as close to the Earth as Marks’ scale model suggests. Moreover, the only drawing made by Mark of the Earth-Moon system illustrates an evenly adequate scaling (see Fig. 4). We interpret that Mark used his mental model of the Earth-Moon system, illustrated by the scale model on the table, to formulate a logical and coherent explanation for the variation in visual size of the Moon.

Moreover, Mark actively rejects the reasoning of Fig. 4 as a cause for variation in apparent Sun size. As stated earlier, he argues that it has something to do with the much larger distance between the Sun and the Earth (see the statement of line 136 above). Based on this quote, we hypothesize that Mark realizes that the relevance of the geographic position effect depends on the ratio of the Earth’s size to the distance to the observed object. Since this ratio is much smaller for the Sun than for the Moon, Mark only judges it relevant for variation in apparent Moon size and not in apparent Sun size. Again, this reasoning is very logical and much in agreement with the scale model on the table—in the sense that the model Sun is much further from Earth than the model Moon, so that the effect of moving one Earth radius would indeed be much smaller for the Sun. While parts of his explanations may be created on the spot (as suggested by the occasional high frequency of hesitations, for instance in lines 56 and 58), we interpreted this as Mark trying to make his reasoning as internally consistent as possible. He attempts to make plausible statements with solid explanatory power, which are coherent with the posed scale model.

When comparing the apparent sizes of the Moon and the Sun, Mark seems to understand that a solar eclipse could not occur with the two discs he chose earlier, because the Sun disc was larger than the Moon disc. He therefore reasons that on such exceptional occasions, they must be of equal size (line 146). This is interpreted as a cognitive conflict: an occasion where a previously held belief (the Sun looks larger in the sky than the Moon) is at odds with newly considered information (during solar eclipses the Moon and the Sun must be of similar apparent size). As a result, there is an attempt to slightly modify the previous conception (the Sun is *usually* larger than the Moon, but there are exceptions) to integrate the new information into

the mental model. This reaction neatly illustrates the malleability of such student conceptions and their susceptibility to changing contexts (moving from the context of apparent sizes to the context of solar eclipses).

Although Mark could correctly explain the principle of a solar eclipse, he did not (adequately) use the scale model to evaluate the plausibility of his explanation. More specifically, he did not critically evaluate the much smaller angle subtended by the tiny Moon compared to the large Sun. This contrasts with his earlier reasoning about the variations in apparent sizes, where he was able to infer from the involved distances that a variation in apparent size with different positions of the observers was more plausible for the Moon than for the Sun. It is, however, very similar to the discrepancy between the visual sizes of the Moon and the Sun as indicated by the chosen discs and by the scale model (because the apparent sizes of both bodies were much larger as shown by the scale model than by the discs). It can therefore be hypothesized that Mark tries to use his understanding of the involved scales to come up with reasonable explanations, but that these attempts fail when the posed problems require mentally taking the perspective of the scale model Earth and imagining certain subtended angles.

When asked about the difference between an annular and a total solar eclipse, Mark referred again to his earlier argument of geographical location dependence (Fig. 5). This suggests how, once formulated, this idea had fixated itself with Mark and was raised to explain varying phenomena while maintaining an overarching consistency.

G. Conclusion—FWT

Our analysis of Mark's interview can be summarized in a few key take-away points. First, Mark tries to provide explanations that are coherent with his scale model of the Earth-Moon-Sun system. Some answers seemed to be derived by considering relevant features of the scale model. For example, the different apparent size of the Moon for different geographic locations (Fig. 5) may be based on a critical evaluation of the distances between the equator and Moon and between the pole and Moon in the scale model. The fact that Mark judged this effect to be irrelevant for the apparent size of the Sun further adds to this interpretation. This search for coherence is much in line with the framework theory point of view. However, taking into account the scale model while explaining phenomena appeared to be too difficult when it required imagining the angles subtended by the Moon and Sun.

Second, the answers provided by Mark show a high degree of consistency over several questions. For example, when explaining the difference between an annular and a total solar eclipse, Mark referred back to his previous argument involving different geographical positions. Mark occasionally also attempted to reconcile two contradicting answers to recover consistency. For instance, the event of a

total solar eclipse was contradictory to his previous answer that the Sun always looks larger than the Moon. Mark recuperated consistency by stating that usually the Sun appears larger than the Moon, but in the exceptional case of a solar eclipse, they can be of equal size. Again, we believe that this adds to the framework theory account.

Third, some of Mark's explanations could be classified as synthetic models, consisting of both scientifically correct elements and naïve conceptions. For example, his idea of the Earth's orbit around the Sun contains not only the correct element of ellipticity but also an incorrect overestimation of the eccentricity, leading to a noticeable difference in apparent Sun size between aphelion and perihelion. This is a known misconception that leads to an incorrect understanding of the seasons (see, e.g., [55]), which is also the case for Mark.

In conclusion, we recognize Mark's search for consistency, both between several answers and between his answers and the scale model on the table, aligning with a core feature of the framework theory. The finding that Mark often succeeds in maintaining consistency suggests that he has a coherent mental model of the Sun-Earth-Moon system. We described several instances where Mark seemed to use his mental model to explain the discussed astronomical phenomena.

VII. MINI-PAPER 2: KNOWLEDGE IN PIECES AND THE RESOURCES THEORY

Author: Lauren Barth-Cohen

A. Intro—KiP&R

The overlapping theories of knowledge in pieces and resource theory started with a focus on students' reasoning about forces and other topics within mechanics [5,56,57]. In the years since, these theories have been widely applied to empirical studies of physics learning and knowledge within topics associated with mechanics [58–61] and have been applied elsewhere to a range of related topics in chemistry and biology that emphasize mechanical processes [62,63]. Across these studies, the focus is often on the structures of one's knowledge system, changes in that system over time (learning), and the underlying learning mechanism. Results of these studies have contributed new knowledge about learning processes with implications for physics teaching and curriculum design [27,64]. But yet, there is only a small and fragmented body of work applying knowledge in pieces and resource theory to topics within astronomy education [23,65–67]. One potential reason for this limited emphasis might be the focus on reasoning about relative distances and sizes in astronomy, which is a different reasoning challenge than forces. Thus, it becomes an open question: To what extent can a theory of learning that has previously focused on learning about mechanics be applied to conceptualize learning about

relative distances and sizes? To begin addressing this question, here we examine interview data in which one student reasoned about the apparent size of the Moon as a single case to begin expanding this theory to topics of relative distances and sizes in astronomy.

B. Theoretical background: Resource theory and knowledge-in-pieces

To conceptualize a learner's knowledge system, I use the overlapping theoretical models of resource theory [56] and knowledge-in-pieces [5]. These models view one's knowledge system as a complex conceptual ecology of various pieces of knowledge, often called resources, that can be productively activated depending on the sensemaking needs at a moment in time. These resources are often referred to as "small reusable pieces of thought that form concepts and arguments" [68]. Resources operate across grain size levels from small intuitions to larger conceptual structures, along with beliefs and concepts. These pieces are built on a variety of everyday and formal learning experiences. All these resources have the potential to be productive for the learner, depending on the context in which they are applied.

One type of piece to be discussed below is p-prims. These are small-scale, intuitive pieces of knowledge about the physical world. p-prims can provide learners with a sense of naturalness and the ability to predict some events or account for feelings of surprise at how some unexpected events come about [5,69]. They are "primitive" in the sense that when they are cued, no further explanation feels needed. In the data, this might appear as the learner having a sense of "this is just how things are." Additionally, p-prims are sensitive to the local context, and they are often evoked as a whole.

Over time, a p-prim might become more or less likely to be activated (or cued) in a given context. There can also be increases or decreases in the relative strength of the connections, and new pieces can emerge and then become more or less applicable in a context over time. Additionally, there can be increases or decreases in the relative strength of the connections; collectively, this results in shifts, as some groups of pieces come to have tighter connections while other groups of pieces might shift to have weaker connections. More broadly, this whole process results in a more coherently structured network over time as one develops more expertise in an area.

C. Aim—KiP&R

Resource theory and KiP have been widely applied to many topics in the introductory physics curriculum, including those that emphasize forces, for instance, fluid dynamics [61]. However, it has not been applied to a focused analysis on the astronomical scale. Given that reasoning about astronomical scale centers on comparisons of relative sizes rather than forces, here we aim to

examine to what extent this theory can be applied to conceptualize learners' reasoning about astronomical scale and what insights it might produce.

D. Method of analysis—KiP&R

To conduct the resource theory analysis, I followed the analysis steps described by the observe, schematize, and systematize (OSS) cycle for identifying knowledge resources [70]. Aligned with this process, I started by watching the video while writing sequential content log notes about the interviewee's reasoning over time, with a focus on the content within each talk turn. From these notes, I then identified moments in time of deeper interest. I iteratively watched the video multiple times while generating a more detailed transcript. At this point, I created more detailed analysis notes and questions at the level of talk turns associated with individual lines of the transcript as the unit of analysis. Sometimes a turn of talk included multiple sentences or a single word or phrase, while sometimes a turn of talk included multiple ideas, and other times a single idea might span multiple turns of talk. I would often watch and rewatch short clips many times to generate more detailed notes and multiple perspectives. Beyond working bottom up from the data, I was also considering existing scholarship on p-prims [5] while working to potentially identify existing p-prims in the data. At that point, I built an argument for the existence of certain p-prims in data, based on the existing scholarship and my detailed analytical notes from the transcript. During this period, I worked across the data and considered similarities and differences in the data across different points during the interview. I also focused on building competing arguments and multiple interpretations as I worked toward conclusions.

E. Results—KiP&R: Two factors impact the apparent Moon size: One's location on the Earth and the Moon's orbit

In this first sequence, we see Mark generating an explanation for two factors impacting the Moon's apparent size. First, the Moon appears differently at two locations on Earth, when one is standing at the equator versus when one is located in Sweden. Second, the elliptical orbit of the Moon impacts the Moon's apparent size as it gets closer and farther from the Earth over time. The episode begins with the interviewer asking Mark to complete a task that involves comparing paper disks of the Moon with the Moon as we see it in the sky. Notice that when choosing a paper disk, Mark hesitates (line 48), and the interviewer asks Mark if he understands the question (line 49), which suggests some uncertainty or, in the moment, figuring it out. After some back and forth, the interviewer winds up asking Mark: "Do we perceive it [the Moon] differently depending on where on Earth we're standing or..?"

Mark answers in the affirmative: “Yeah, uh, well I mean here in Sweden, we’re kind of close to, you know, the top of the Earth so I, yeah, I’d say so, yeah” (line 54). A moment later, he continues expanding this idea with the notion that the top of the Earth is farther from the Moon, compared to other parts of the Earth, and thus the apparent size of the Moon would change (line 56). For this statement, it is important to notice both the content and the linguistic dimensions of the speech. There are several pauses, hesitations, and interjections (e.g., “well I mean”; “you know”; “yeah, I’d say so, yeah”; “Yeah... oh wait, actually”; “you know it’s... It’s a bit more, you know”) as Mark explains these ideas. His use of these phrases suggests that he is thinking about the ideas in the moment as he constructs an explanation, rather than drawing on a previously constructed or deeply held explanation. There are many reasons why individuals pause, hesitate, and use interjections, for example, nervousness or fear of saying the incorrect thing. However, in this episode, when we compare these lines of transcript with other earlier or later interview moments, these pauses, hesitations, and interjections are less pronounced. That difference suggests that during this interview, Mark is not systemically or uniformly speaking this way; instead, his interjections are more pronounced at this moment, when explaining why one would be perceiving the Moon differently at different locations on Earth. Thus, Mark is likely in the moment constructing an explanation about the Moon appearing different when one is standing at the equator versus in Sweden.

The conversation continues with the interviewer explicitly asking Mark about how the size of the Moon would change on the equator as compared with the top of the Earth (line 59). Mark answers “yeah,” and then the interviewer asks how different it is and points at the Moon disks (line 61). Mark mentions an incremental change (line 62), and then he continues building on that idea when again, restating that the Moon appears bigger when standing on the equator and smaller in Sweden (line 64). While again explaining this idea, Mark physically picks up the relevant Moon disks with his hands, and at this moment he might be a tad more comfortable with the explanation, given fewer interjections, than previously after explaining the same line of reasoning multiple times, he is gaining comfort with his explanation about the Moon appearing different at different locations on Earth.

Then there is a switch in the conversation as they focus on discussing how the orbit of the Moon impacts the Moon’s appearance. The interviewer asks a follow-up question about observing the Moon at different times in Sweden and what variation might be seen (line 65). With this question, there is a conversation shift. Mark discusses the Moon going further or closer to you when in an orbit, but that the change would not be as noticeable as the change from watching the Moon at the equator (line 66).

The second sequence begins with the interviewer asking about whether there is a difference when staying in Sweden and how this relates to picking a range of disks or if one disk is always the best (line 67). In answering this question, Mark says that it is always this one disk (line 68). The interviewer then asks him to go further on the “why” as related to the orbit (line 69), and then Mark brings up the idea of the Moon’s orbit not being a complete (perfect) circle and being an ellipse instead (line 70). The accompanying gesture of moving hands closer and farther seems to support the notion of an ellipse. Also, notice that Mark’s hesitation and interjections (line 70) might be suggestive of him figuring out his idea in the moment, pausing and first saying “yeah” then saying “wait no” and finally appearing to overall agree with the interviewer’s prior suggestion that it has to do with an orbit. The interviewer then asks for more information, and his follow-up explanation is consistent with the idea of the Moon’s orbit as elliptical, with some locations being closer or farther from the Earth at the center (line 72). Similar to before, notice some hesitation and uncertainty given the phrases (“Euh, I think” and “But I’m not sure obviously”). Mark is explicitly unsure if the path of the Moon is a complete circle or not. Similarly, this suggests that he might be thinking about these ideas in the moment, as he is unsure while constructing an explanation. While speaking the phrase about being “unsure,” Mark tugs on his shirt at the elbow in a manner that potentially suggests some uncertainty or hesitation with the ideas. During the final portion of this line, Mark gestures to show that the distance between him and the Moon would be shorter or longer at different points in the orbit. The final portion of this line, with the accompanying gesture, does not have the same sense of unsureness as the earlier parts of this transcript line, given the lack of the prior phrases related to hesitation. Next, Mark generates a drawing, and the interviewer summarizes that there are two factors (orbit of the Moon and position of the observer) impacting it all (line 85), and Mark agrees (line 86).

Following that moment, the interviewer asks Mark about his memory of times in the past when he looked at the Moon and noticed the relative size (line 87). In responding to that question, Mark says “no.” That is, he explains that he does not look at the Moon much, and the thing he has most noticed is the shading and changes in phases over the month (“Euh, not, not really. I honestly, I don’t look at the Moon that much, I mean it’s mostly... the thing I noticed about the Moon is mostly just how one part of it is shaded and one part of it’s, you know, if it’s a half Moon, if it’s a full Moon or, you know” line 88). A moment later, Mark further reiterates that he has never thought about size (line 92). This is important information as one considers how much analytical weight to give Mark’s interpretation of the interview questions that asked him to evaluate various disk sizes. Given that Mark

explicitly says he does not notice the Moon sizes, the comments suggest he is answering these questions while thinking about something new to him, the size of the Moon, and is not explaining a concept that he had previously considered deeply. Therefore, from the analytical interpretation perspective, we should not emphasize the specific size too much since he says he does not notice it.

To summarize, across these two sequences, Mark is discussing two factors that impact the apparent size of the Moon: the location one is standing on Earth and the Moon's orbit around the Earth. Regarding the first factor, Mark argues that the Moon looks bigger when standing on the equator and smaller in Sweden. Based on Mark's hesitation and interjections, along with his explicit statement that he does not look at the Moon very much, I argued that he is constructing these explanations in the moment. As an important additional consideration, from the scientific perspective, the Moon is so far from Earth that one's location on Earth does not result in a large enough change to impact the apparent size of the Moon from the Earth. That is, there is a tiny size difference, but it would be hard or impossible for an observer to detect this difference. The Moon does vary in how far it is from the Earth over the month, and that variation results in about a 10% change in the Moon's apparent size. Given the elliptical orbit of the Moon, it can be noticeably bigger at times. As Mark explicitly says that he "never really [thought] about the size," which suggests that he has not noticed this issue and has not thought about it before even if it is noticeable on Earth, that is further evidence that Mark has not detected differences in the apparent Moon size.

Next, I consider what knowledge pieces or resources Mark might be activating while generating this explanation. From the perspective of knowledge in pieces, Mark's explanation for the changing size of the Moon as due to different locations on the Earth is potentially drawing Ohm's p-prim ("more is more"). This intuitive idea is that more of one item corresponds to more of another item; that is, for example, being closer to a heat source contributes to feeling hotter [5] or, as another example, more sand particles joining a sand dune, resulting in a larger dune [71]. Elsewhere, research has documented this p-prim as applied to learners' explanations for the seasons, that is, the Earth being closer to the sun in the summer and farther in the winter (closer = hotter) [56,72]. Differently, in this episode, Mark is applying this knowledge piece to his explanation for the changes in the Moon's size. That is, being closer to the Moon (on the equator) contributes to the Moon being larger, and being farther from the Moon (in Sweden) contributes to the Moon being smaller (closer = bigger).

When discussing the second factor impacting the apparent size of the Moon, the orbit, Mark does not explicitly state how the orbit impacts the apparent size (although this idea is discussed more later during the

interview). The lack of a causal mechanism for this explanation can be viewed in any number of ways, possibly as evidence that he constructed an explanation in the moment without deeper considerations, possibly because he did not view a causal mechanism as important to mention, or possibly due to the specifics of the interview social dynamics, which made it not relevant at the moment. Another possibility is that if different modes of communication (as in semiotic resources) were utilized, they might have impacted his reasoning.

When comparing these two factors, the location on the Earth and the Moon's orbit, Mark mentions that the orbit is not as important for impacting the size of the Moon as the location on the Earth. This is an interesting line of reasoning and raises questions regarding various pieces of knowledge and different priorities. Recall from the above discussion that Mark applied Ohm's p-prim to the changing size of the Moon due to different locations on the Earth. Given that Mark used Ohm's p-prim for this line of reasoning, we can ask why he didn't use this p-prim elsewhere, where it could have hypothetically been used. Specifically, Mark did not use Ohm's p-prim to reason about how when the Moon is closer to the Earth, it appears larger, and when it is farther away, it appears smaller. Perhaps for some unknown reason, Mark views this p-prim as not applicable to this context. Another possibility is that he did use Ohm's p-prim in this context of the closer orbit and the Moon appearing larger, but did not state his ideas aloud for any number of reasons. Thus, this work is potentially documenting the context-specificity of this p-prim being verbally activated. Recall that at this moment in the interview, Mark argued that the location where one is standing on the Earth is more important than the distance to the Moon. This raises questions about why Mark might view the location on Earth as more important at this momentary line of reasoning. Potentially, something in his life experience, living in Sweden relatively close to the pole, cues him to view the location on Earth as an important factor in this line of reasoning.

F. Discussion—KiP&R

Across these two sequences, there is some potential evidence for Mark having constructed an explanation on the spot, given various uncertainties and hesitations. Importantly, later during the interview, he explicitly mentions that he had never really considered the Moon's size, which is stronger evidence of him constructing an explanation on the spot, rather than drawing on a preexisting explanation. Knowledge underlying this explanation includes a combination of Ohm's p-prim to account for a closer Moon appearing larger, along with potentially something related to his life experience, living in Sweden relatively close to the pole, which cues him to view the location on Earth as an important factor. Interestingly, when juxtaposing these two explanations

for the Moon's size, Mark argued that the location where one is standing on the Earth is more important, but it is unclear why he views it so. It might involve issues of relative distances and ratios or a question about which knowledge pieces have higher or lower priority.

From the perspective of knowledge in pieces, the results of this analysis document an instance of a learner constructing an explanation on the spot based on a collection of knowledge, including Ohm's p-prim and life experiences in Sweden. Across the interview, there might be increasing comfort with the explanation, which is to be expected with the application of similar pieces multiple times. Importantly, this analysis touches upon a larger question about how different life experiences, potentially across different cultures, impact knowledge, as has been extensively discussed elsewhere [73,74]. Here, we contribute an evidentiary example in which knowledge rooted in geographically specific life experiences is influencing reasoning about a physical phenomenon. Another contribution is that this analysis captures an instance of a p-prim that has primarily been documented within the context of force and motion being applied to astronomical distances. This is interesting within the larger question of the range of applicability of knowledge in pieces to new topics [5].

G. Conclusion and implications—KiP&R

When considering applications to astronomy teaching, learning, and instruction, this episode highlights the complexities involved in learners' considerations of the apparent size of astronomical objects. That is, this case highlights that learners are not necessarily aware of the size of objects seen in the sky and their relative size to other objects. Astronomical distances are challenging for learners, including relative distances, even when they bring relevant, intuitive, and life experiences to consider the phenomena. More specifically, we saw that a learner can struggle to identify what factors impact our perceptions of astronomical information. The analysis documented how the learner brought existing knowledge to construct explanations about these topics and the various struggles encountered. Implications suggest the need for care when designing and implementing instruction on these topics, given the complexities highlighted.

VIII. MINI-PAPER 3: SOCIAL SEMIOTICS (SS)

Authors: Christopher Robin Samuelsson and Urban Eriksson

A. Intro—SS

Learning physics requires students to engage with and generate multiple forms of representation: gestures, drawings, mathematical notation, spoken and written language, etc., and to coordinate meaning across these forms [75–77]. Within social semiotics, these forms are

understood as semiotic systems, and their coordination constitutes what is broadly termed representational competence. Such competence is particularly important in astronomy education, where the objects of study generally cannot be directly seen or manipulated and where reasoning necessarily involves spatial relationships between entities at vastly different scales. A key process in the development of representational competence is transduction: the movement of meaning from one semiotic system to another, for example, from gesture to drawing or from speech to mathematical notation [76,78]. Research has established that multimodal approaches promote learning, although the microlevel dynamics of how students enact transductions, the moment-by-moment shifts between, say, an ephemeral gesture and a persistent drawing, remain less well understood [79]. When students spontaneously shift between semiotic systems, they engage in real-time meaning-making that can reveal how disciplinary understanding develops. Attending to these transductive moments provides insight not only into what students understand but also into how representational competence is actively constructed. The present study addresses this gap by examining how a high school student, Mark, transduces meaning between speech, gesture, and drawing while reasoning about the apparent size of the Moon from different observational positions on Earth. By applying social semiotic analysis to Mark's engagement with an activity exploring the Earth-Moon system, we trace how meaning is carried across semiotic systems and what this reveals about the development of disciplinary understanding.

B. Theoretical background: Social semiotics

Social semiotics provides a comprehensive theoretical lens for understanding how meaning is constructed and communicated within specific communities of practice. In educational contexts, social semiotics can be defined as “the study of the development and reproduction of specialized systems of meaning-making in particular sections of society” [11] (p. 95). Physics classrooms constitute one such section of society, with their own specialized systems of meaning-making that students must learn to navigate. In this context, learning involves not merely acquiring factual knowledge but developing facility with the representational practices through which physics knowledge is constructed and communicated. This approach recognizes that learning, or meaning-making, involves coordinating multiple modes of communication beyond just verbal language. Within the model, we refer to “those aspects of physics concepts that have particular relevance for carrying out a specific task” [75] (p. 2) as disciplinary-relevant aspects (DRA), for example, when discussing the size of the Moon, the DRA's include the distance between Earth and Moon.

Central to social semiotics are the concepts of semiotic systems and semiotic resources—the tools and modes through which meaning is made and shared within a community. In educational settings, examples of commonly used semiotic systems and resources include “graphs, diagrams, sketches, figures, mathematics, specialist language” [11] (p. 96). Unlike static representations, semiotic resources are understood to have meaning potential that is realized through their use in specific contexts [11]. A semiotic system can be thought of as a category of semiotic resources (e.g., color—red, blue, green, or gestures—symbiotic and representational).

When a small group of physics students is working on a homework problem together, they might make meaning through multiple semiotic systems, including verbal language, drawings, and equations, and further share that meaning with their larger physics class community through a whiteboard presentation. Each student may individually transduce the meaning between different semiotic systems; one may begin by talking about their reasoning and then move on to formulating that reasoning with mathematical notation, and another student may begin by drawing a figure to structure their reasoning before moving on to sharing it verbally with the others. Finally, the students may negotiate about what they can agree on and formulate it on a whiteboard. All these shifts between representations and ways of expression involve semiotic systems and transductions.

1. Persistent and nonpersistent semiotic resources

The social semiotic model makes a crucial distinction between persistent and nonpersistent semiotic resources. Persistent representations include stable, enduring forms such as different types of graphs, drawings, and diagrams, while nonpersistent representations encompass transient forms like types of speech and gesture. This distinction is particularly relevant for understanding how different semiotic resources coordinate in educational interactions.

Research suggests that persistent representations often can act as hubs around which nonpersistent representations can be coordinated [80]. This creates a hierarchical relationship where gestures and speech orient toward and help interpret more stable visual or material resources.

2. Transduction and multimodal communication

A key concept for analyzing meaning-making within social semiotics is transduction—the process by which meaning is carried across different semiotic systems. A central analytical distinction in this model is between transduction and transformation. Transduction refers to the process of carrying meaning across different semiotic systems, for example, when a student expresses through a drawing what was previously communicated through gesture. Transformation, by contrast, involves changes in meaning within the same semiotic system, for example,

when a student modifies a drawing by adding new elements to it [76]. This distinction is analytically important because transduction, as a cross-system process, places greater demands on the meaning-maker: the affordances of the target system may foreground aspects that were backgrounded in the source system or vice versa. It is primarily through transduction that we can observe how students’ understanding is reshaped as they move between representational modes. Gregorcic *et al.* [81] illustrate how students “juxtaposed talk with gestures and resources in the environment to communicate ideas that they initially were not able to express using words alone” (p. 1). This suggests that a system like gestures often may serve as transductive bridges, helping to move meaning between verbal, visual, and spatial modes.

3. Gesture as a semiotic system

Among the various semiotic systems available to students, gesture plays a particularly important role in physics reasoning. Gregorcic *et al.* [81] demonstrate how students “relied heavily on nonverbal meaning-making resources, most notably hand gestures and resources in the surrounding environment” (p. 1) when engaging with complex physics concepts. Gestures function not merely as accompaniments to speech but as integral components of meaning-making that can “express relatively fluently complex ideas on a novel physics topic” (p. 1). For instance, when explaining electromagnetic induction, a teacher who moves her hand through an imaginary loop while saying “the magnet passes through the coil” produces a representational gesture. In contrast, a quick downward hand chop used to emphasize the word “induced” constitutes a beat that signals importance without representing any physics content.

Research distinguishes between different types of gestures based on their meaning-making functions. Representational gestures carry explicit meaning related to the disciplinary content, while nonrepresentational gestures (such as beats) serve rhythmic or emphatic functions without conveying specific conceptual content. Gregorcic *et al.* [81] note that representational gestures are “the most interesting category of gesture for the physics education research community” (p. 3) because it is a category of representation “explicitly relating to physics content” (p. 3).

Of particular importance for educational analysis are symbiotic gestures—hand movements that “are produced against an external, perceivable structure in the environment, and rely on the meaning provided by the environment to make full sense” [81] (p. 2). These gestures derive their meaning through coupling with environmental features, such as diagrams, digital displays, or physical apparatus—for instance, holding one’s arm at the same angle as the sloped line in a graph. A symbiotic gesture could also be representational if the gesture itself carries

meaning while being produced or carried out together with a part of the environment, like in the previous example: representing a slope with the arm against the sloped line in a graph gesturing (the gesture and the line in the graph carry similar meaning) or gesturing two drum sticks with two fingers drumming a drum set that has been drawn on a piece of paper (the gesture and the drawing have different but supplementary meaning). However, a gesture that points at something in a picture is symbiotic but not representational. The model recognizes that “none of these systems in isolation would be sufficient to construct the actions that the participants are pursuing” [82] (p. 36), emphasizing the importance of analyzing the “ecology of sign systems” (p. 36) rather than individual semiotic resources or systems in isolation. For gesture analysis, this means understanding how hand movements coordinate with speech, environmental features, and other meaning-making resources. We refer to these types of “ecologies” or coordinated semiotic resources from different semiotic systems as systems of semiotic resources.

This social semiotic approach provides a comprehensive framework for understanding student gestures not as mere behavioral phenomena, but as integral components of disciplinary meaning-making processes that coordinate with other semiotic resources to enable learning and communication.

C. Aim—SS

This study investigates how a high school physics student generates and coordinates different semiotic resources when reasoning about the Earth-Moon system. Specifically, we examine which semiotic resources Mark engages with and how meaning is transduced and translated between different representational modes. Understanding these transductive processes is important because students’ personal transductions provide insight into what and how learning is taking place [76]. Such insights can inform how physics teachers might better support students in developing representational competence. Additionally, attending to these microlevel transductions offers a window into meaning-making processes that would remain invisible if analysis focused only on final representational products rather than the dynamic shifts between semiotic systems through which representational competence develops.

D. Method of analysis—SS

1. Grain size

Our analysis operates at the grain size of individual representations and equipment interactions. We examine each piece of representation or equipment that Mark engages with while reasoning about six DRAs: orbit of Moon, Moon, Earth, position of Moon relative to Earth, distance between Earth and Moon, and orbital ellipticity

(“Bulgyness,” in Mark’s words). This can be one type of unique gesture, a drawing of an object (Earth) together with the spoken lines of reasoning accompanying the nonspoken semiotic resource.

2. Elements/aspects of interest

In the analysis, we are mainly interested in the semiotic resources and systems (interacted or generated representations and equipment) and how they are coordinated or sequenced together. In particular, we are interested in the representational and nonrepresentational gestures engaged in when reasoning about the Earth-Moon system for the particular DRAs described in the grain size section.

3. Transcription

Video data of Mark’s engagement with the activity were collected and subsequently processed through multimodal transcription. The transcription process categorized the interactions into three distinct but interrelated components:

Speech: This category captured the verbal exchanges between the researcher and Mark, including questions, responses, hesitations, and verbal reasoning. The transcription preserved linguistic features such as pauses, emphasis, and uncertainty markers (e.g., “Uh...well, I mean,” “I’d say like maybe”).

Gestures: All nonverbal communicative movements, with the hands, were documented in this category, with particular attention to the typology of gestures described in the model (representational and nonrepresentational gestures).

Interaction: This category documented all instances where Mark engaged with persistent semiotic resources in the environment, including creating something (drawing a picture) and manipulating objects (like moving the sticks with the Moon).

The transcripts were integrated directly into the video stream, creating a synchronized multimodal record where each category could be precisely matched to the others and to the corresponding video data. This integration enabled a moment-by-moment analysis of how different semiotic resources were coordinated in real time. For this purpose, we used ELAN—a video annotation tool developed by The Max Planck Institute for Psycholinguistics [83].

4. Analysis

Our social semiotic analysis focuses on two key elements: the semiotic resources Mark uses for meaning-making, and the transduction processes through which meaning is carried between different representational modes, i.e., transduction [11,76]. Within this model, we define a semiotic resource as any type of tool for meaning-making (e.g., symbols, equipment, body language, or speech) that participants engage with or discuss.

A system of semiotic resources may not necessarily constitute a semiotic system, as the resources may belong to different semiotic systems. In our analysis, we describe two types of gestures that have been described in earlier research on gestures in science, technology, engineering, and mathematics education [81,84,85]:

- Representational (REP): This includes gestures that convey meaning, for example iconic gestures through motion or shapes, e.g., Mark's hand movements forming a sphere to represent the Moon [84].
- Nonrepresentational (NONREP): This includes rhythmic or pointing gestures that do not convey any meaning in themselves but that may be carried out together with other semiotic systems that convey meaning, sometimes to accentuate aspects of another semiotic system, e.g., rhythmic beating gestures that accompanied speech [81,85].

The analysis focuses on Mark's engagement with multiple semiotic systems while reasoning about the apparent size of the Moon as observed from different positions on Earth and at different times. The semiotic systems involved include:

- Gestures (nonpersistent)
- Speech (nonpersistent)
- Equipment (persistent): Moon and Sun on sticks, pens
- Drawings (persistent)

The last two systems are usually involved in the interaction component, gestures in the gesture component, and speech in the speech component of the transcript. The analysis examines how Mark coordinates these different semiotic systems to communicate ideas and make decisions, with particular attention to instances of transduction between systems. This is important for getting insights into shifts in meaning-making in the conversation. Quotes from Mark are added in the text with line numbers. Note that these quotes are slightly different from those in the shared results (Sec. V), since SS requires a different transcription, including e.g., rhythmic gestures.

E. Results—SS

Mark's engagement with the activity reveals distinct patterns in his use of semiotic resources. Generally, he engages with persistent semiotic resources when making decisions without hesitation. For example, when asked about the difference in apparent size of the Moon when observed from two different positions on Earth ("equatorial" and "Swedish" position), Mark promptly utilizes physical objects, stating: "I'd say like maybe just like...small incremental change maybe, maybe like this [holds up two sticks with images of the Moon of different sizes]." (line 62).

Two distinct sequences were identified that involve different patterns of semiotic resource engagement. Both sequences involve a limited number of DRAs [75] related to the Moon-Earth system: the Moon, Earth, the orbit of

the Moon, distance between Moon and Earth, and positions of the Moon (relative to Earth). The sequences are described in the next two sections and highlight the different levels of transduction.

1. Sequence 1: Limited transduction with primarily symbiotic gestures

This sequence begins with a question about whether a selected image of the Moon (on a stick) represents the apparent size that a person would observe from a specific position on Earth.

When responding to questions about the apparent size of the Moon from different positions on Earth, Mark primarily uses pointing gestures that reference elements in the drawing:

Mark: "It's just...it's just like if you're standing here [points at postit-note], is this how it's going to look [points at postit-note and makes a motoric rhythmic gesture]?" (line 52).

These gestures are nonrepresentational, referring to the drawing rather than conveying meaning independently. There is minimal transduction between semiotic systems in this sequence. Mark continues:

Mark: "Uh...well I mean....here in Sweden [points at the postit-note], we're kind of close to, you know, the top of the Earth [holds out hand] so I [points at the postit-note], yeah, I'd say so [points at the postit-note]." (line 54).

Again, the gestures are nonrepresentational. When asked whether the apparent size of the Moon varies at different times rather than different positions, Mark relies exclusively on rhythmic gestures:

Mark: "Uh...well, I mean it would technically go further away or closer to [beating rhythmic gestures], you know, where I'm standing [beating rhythmic gestures], you know, when in its orbit [sweeping rhythmic gesture]. Uh...so...yeah, I think it might you know also like change a little bit [beating rhythmic gesture], but not, not as noticeably [beating rhythmic gesture] as eh... to like watch it at the equator compared to ["chopping" rhythmic gesture]... yeah." (line 66).

Rhythmic gestures are primarily used to accentuate arguments and are, like the pointing gestures, nonrepresentational. Additionally, in this portion of the sequence, Mark employs primarily nonpersistent semiotic resources (speech and rhythmic gestures) that are not transduced into new semiotic systems. The limited transduction observed in this sequence is itself analytically significant. Mark's

reliance on nonrepresentational gestures and his tendency to remain within the speech-and-pointing mode suggest that the question as posed does not generate sufficient conceptual friction to prompt him to seek alternative representational resources. His reasoning stays close to the surface of the persistent drawing without engaging with it generatively. It is worth noting that this absence of transduction is not merely incidental but structurally constrained. Nonrepresentational gestures, by definition, do not carry disciplinary meaning in themselves, they serve rhythmic or emphatic functions in coordination with other semiotic systems, most typically speech. Since transduction involves the remaking of meaning across semiotic systems [76], a gesture that carries no meaning cannot serve as either a source or a target for transduction. Mark's predominant use of rhythmic gestures in sequence 1 therefore not only correlates with limited transduction but effectively precludes it. An analogy may be helpful: just as a musician might hum along to a familiar tune without needing to consult the score, Mark can address the question about apparent size by referencing the drawing without needing to reconstruct or extend the system of representations. It is only when the conceptual demands increase, as they do in sequence 2, that the limitations of this mode become apparent and transduction becomes necessary.

2. Sequence 2: Complex transduction between multiple semiotic systems

The second sequence begins when Mark is asked how the shape of the Moon's orbit relates to his reasoning. Initially, he engages only with nonpersistent resources (speech and gestures):

Mark: "Euh...yeah... oh wait no,... [holds hand under chin] I mean it depends on [motoric rhythmic gesture] like if the Moon orbit is like a complete circle [circling gesture with one hand] or if, you know, it gets shorter [moves hands closer to each other] in some spans or longer [moves hands away from each other] in some spans. [crosses his arms]" (line 70).

Here, Mark engages in representational gestures in order to reason about the orbit of the Moon (circling hand) and the variation in distance between the Moon and Earth (shifting hands closer and further away from each other). Mark continues:

Mark: "[with arms crossed] Euh...I think I think it isn't a complete circle eh.... But I'm not sure obviously [motoric rhythmic gesture] eh... So, you know, sometimes when eh the Moon is around [shapes a sphere with his hands that he then moves in an orbit around his own body] the span between

me and the Moon [moves right hand between the "hemisphere" formed by his left hand and his chest to indicate distance] would then be shorter than other parts of [repeats the gesturing but moves it a bit to his left to indicate a new position on his body in relation to the "hemisphere" indicated by his hand], you know, between...[motoric rhythmic gesturing]" (line 72)

Here, Mark again engages in representational gestures of some of the DRA's previously mentioned. Mark engages with a system of nonpersistent semiotic resources to represent multiple complex relationships simultaneously: the Moon's orbit (motion of hand circling Mark's body), the Moon (Mark's hands shaping a sphere) different positions of the Moon over time (left hand moving relative to Mark's body), and the position on Earth relative to each position of the Moon (Mark's body and right hand gesturing between left hand and body). However, maintaining this system while reasoning proves challenging and Mark is asked to make a drawing.

When asked to make a drawing, Mark immediately begins generating a system of persistent semiotic resources that mirrors the structure of his previously established nonpersistent system. This transduction process is evident:

1. He draws the elliptical orbit, the Moon, and Earth
2. He indicates positions with distances drawn as lines between two positions of the Moon in orbit and one (Mark's) position on Earth. Two Moons are drawn—one with a dotted line indicating a shifted position of the Moon and one with a solid line.
3. He adds contextual elements like "E for Earth" and uses a sharp border for one Moon position and a dotted border for another to indicate they represent the same Moon at different positions

This drawing system allows for more specific questions and responses about particular parts of the system, such as which "Moon" position is closer to Earth and why. The persistent nature of the drawing enables Mark to engage more confidently with his reasoning. Mark initially realizes that he is limited in the way he can make meaning about the Earth-Moon system and so transduces some of the meaning through the representational gestures, this meaning is finally, in turn, transduced into the drawing and so a transductive chain is formed [79].

Table I summarizes how each DRA is realized across the three semiotic systems Mark engages with during sequence 2. The table reveals that most DRAs are represented across all three systems, though with notable variation: for instance, the ellipticity of the orbit "bulgyness" appears in speech and drawing but is not represented gesturally, while the use of Mark's own body as a semiotic resource for Earth has no direct analog in the drawing, where Earth is instead represented as a labeled circle.

TABLE I. Disciplinary Relevant Aspects (DRA's) mapped to Mark's gestures and drawings.

DRA	Speech	Gesture (non-persistent and representational)	Drawing (persistent and representational)
Orbit of Moon	"If the Moon orbit is like a complete circle", "sometimes when eh the Moon is around"	Circling hand	Drawn elliptic orbit
Moon	"Sometimes when eh the Moon is around"	Hands shaping a sphere or hemisphere	Circle representing the Moon on the drawing of the orbit
Earth	"The span between me and the Moon"	Body of Mark	Circle in the middle of the system
Position of Moon relative to Earth	"Sometimes when eh the Moon is around"	Motion of hand from one position relative to Mark's body to a new position	One Moon with solid line and one Moon with dotted line
Distance between Earth and Moon	"The span between me and the Moon"	Motion of hand between body and gesturing of the Moon	Lines between Earth circle and Moon circles
Other: Bulgyness of orbit	"It gets shorter in some spans or longer in som spans", "be shorter than other parts of, you know, between"	Hands forming two distances: one shorter and one longer	Drawn "bulgy" shape of orbit

F. Discussion—SS

The analysis of Mark's engagement with multiple semiotic resources reveals several important insights into his meaning-making processes during physics reasoning:

1. Persistent vs nonpersistent semiotic resources and confidence

Mark demonstrates noticeably greater confidence and clarity in his reasoning when engaging with persistent semiotic resources (drawings and physical objects) compared to nonpersistent resources (gestures and speech alone). This heightened confidence becomes particularly evident in sequence 2, where the transition from gestural representations to a concrete drawing enables Mark to articulate more precise relationships between elements of the Earth-Moon system. This supports previous research suggesting that persistent semiotic resources reduce cognitive load by externally maintaining complex relationships [81]. When Mark relies solely on nonpersistent resources or persistent resources provided by the researcher (i.e., that Mark himself has not generated) in sequence 1, his responses contain more hesitations and less specific spatial reasoning.

2. Transductive chain as indication of learning

The process of transduction, i.e., translating meaning from one semiotic system to another, appears to be particularly significant in Mark's reasoning development. The transductive chain formed in sequence 2 represents a critical progression in Mark's meaning-making: from

mental imagery to representational gestures and finally to drawings. This chain of transductions exhibits what Volkwyn [76] suggests may be indicative of learning and deepening conceptual understanding. As meaning is remade across different semiotic systems, Mark appears to refine and clarify his understanding of the discipline-relevant aspects (DRAs) of the Earth-Moon system.

The transductive chain is particularly evident when Mark initially realizes the limitations of verbal explanation, shifts to representational gestures shaping the Moon and its orbit around his body, and finally transduces this meaning into a persistent drawing that captures the spatial relationships more precisely. Each transduction appears to enable more sophisticated reasoning about the relationships between the Earth, Moon, different positions, and distances.

3. Mapping of disciplinary-relevant aspects across semiotic systems

The mapping of DRAs across semiotic systems (Table I) reveals that transduction does not produce identical representations in each system but rather complementary ones. It has been argued that this process "is a sign that learning is taking place" [76] (p. 113). The "bulgyness" of the orbit, for example, emerges only in the drawing, suggesting that the persistent medium afforded a degree of spatial precision that the ephemeral gestures did not. Conversely, the embodied representation of Earth—Mark using his own body as a reference point—captures a first-person spatial perspective that the bird's-eye drawing necessarily loses. This complementarity supports the

argument that different semiotic systems do not merely duplicate meaning but reshape it, foregrounding different DRAs in the process. This distribution of meaning across complementary systems also serves to reduce the cognitive load involved in coordinating multiple spatial relationships simultaneously [86].

Another instance of transduction occurs when these aspects are mapped onto a persistent drawing that Mark makes, where:

- The Moon's orbit is represented as an ellipse.
- Earth is labeled with "E."
- Two positions of the Moon are distinguished (one with solid lines and one with dotted lines).
- Lines are drawn connecting Earth to different Moon positions, explicitly representing distance.

Adding new semiotic systems may also allow Mark to discover or communicate new insights, for example, the "bulgyness" was not represented gesturally but emerges in the drawing. However, it is worth noting a potential tension in the transductive process. Each successive transduction appears to reinforce Mark's commitment to his current line of reasoning: the gesture system mirrors the verbal reasoning, and the drawing system in turn mirrors the gesture system. If each new representational system confirms rather than challenges the previous one, transduction may function not only as a mechanism for learning but also as a mechanism for entrenchment—a kind of representational confirmation bias in which the coherence across systems is taken as evidence for the correctness of the reasoning. This raises a pedagogically important question: at what point might a teacher productively intervene to introduce a representational challenge, for instance, a canonical diagram that differs from the student's own drawing, in order to disrupt an unproductive transductive chain? The relationship between transductive coherence and conceptual accuracy is an area that warrants further investigation.

4. Gesture types and their functions

Mark employs different types of gestures that serve distinct functions throughout his reasoning:

- Rhythmic/motoric gestures primarily emphasize speech but carry limited disciplinary meaning
- Symbiotic gestures reference elements in existing persistent representations
- Representational gestures embody abstract concepts like orbits and distances

Notably, some of Mark's gestures exhibit a duality, being both representational and symbiotic simultaneously—a phenomenon previously observed in students working with planetary orbits [81]. The progression from primarily nonrepresentational gestures in sequence 1 to complex representational gestures in sequence 2 appears to parallel his developing conceptual understanding.

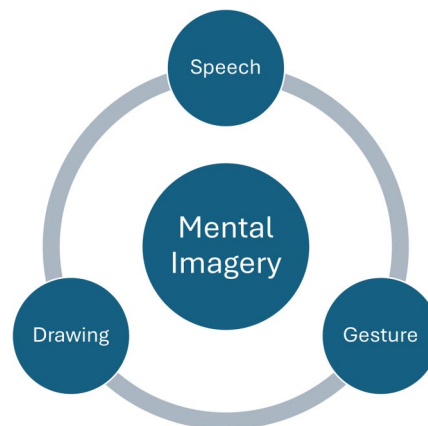


FIG. 6. A transductive chain in which the students' mental imagery is distributed over three semiotic systems through transduction.

G. Conclusion—SS

This analysis demonstrates that the semiotic resources a student engages with fundamentally shape the character of their physics reasoning. Mark's case illustrates two key findings. First, persistent semiotic resources facilitate more complex and confident reasoning than nonpersistent resources alone, supporting the view that externalizing meaning reduces the cognitive demands of coordinating multiple spatial relationships simultaneously. Second, the transductive chain from speech through gesture to drawing (see Fig. 6) constitutes a critical progression in meaning-making: each shift across semiotic systems creates opportunities for the student to refine, extend, and, potentially, entrench their understanding. For physics education, these findings suggest that instructional designs should deliberately scaffold transductive processes, providing students with opportunities, and reasons, to express their reasoning across multiple semiotic systems. Equally important, educators should attend to the transductive process itself as a diagnostic window: the ways in which students carry meaning between systems reveals not just their current understanding but the trajectory of its development. Future research could investigate how targeted representational challenges might be used to productively disrupt transductive chains when they lead toward misconceptions, and how transductive competence develops over longer time-scales in physics learning.

IX. OVERARCHING DISCUSSION

In this work, we analyzed one interview from three different models: framework theory, knowledge-in-pieces, and social semiotics. In this discussion, we elaborate on how these three different perspectives lead to different results, how they can complement each other, and on how the cooperative collaboration leads to new questions.

A. Different models lead to different results

All three researcher teams agreed that Mark's was the most interesting interview (though for different reasons). We also agreed on which part of Mark's interview was most interesting and we agreed that Mark is initially hesitant but eventually builds a more or less coherent understanding.

There are, however, also differences due to the different approaches the three author teams took to analyze and interpret the same data. This is very clear in the research questions or aims (all set within this particular topic).

- **FWT**: look for Mark's mental model and how it is used over various related but different questions.
- **KiP**: To what extent does the KiP approach apply to astronomical scale?
- **SS**: How is meaning transduced and transferred between semiotic resources?

These RQ/aims are different because of different interests, but also because of different epistemological stances. FWT and KiP think it is possible to assess or examine students' (in this case Mark's) thoughts, and to construct an explanation for his mental model or conceptual understanding as seen through the interview question answers that are verbally expressed. FWT does so with an assumption of a more coherent mental model, KiP does not but relies on small elements that come together in a conceptual understanding. SS does not share either of these assumptions, they argue that this is impossible to evaluate a person's thoughts. It is however, as SS argues, possible and useful to look in a detailed way at the way understanding is expressed and shared. Thereby, all three models recognize the importance of verbal expression, but it takes on different roles in the analysis and results.

So, directly comparing, underneath the RQs/Aims are the following different assumptions about learner's knowledge and learning mechanisms.

- **FWT**: Assumption of a mental model and learning as the striving for internal consistency.
- **KiP**: Assumption of a complex knowledge system and learning as the reorganize of knowledge pieces.
- **SS**: Assumption about the importance of representational systems and transduction as central to learning.

These different aims and assumptions about knowledge and learning result in different methodologies. This is expressed in different ways to analyze the data, looking at different grain sizes of analysis.

- **FWT**: employs an approach that integrates all answers, drawings, gestures that follows from a question. The grainsize is thus large, encompassing each question as a grain but through the lens of the whole interview (the largest grainsize possible).
- **KiP**: iteratively analyses utilizing the observe, schematize, and systematize cycle for identifying knowledge resources. The grain size is at the level of a single turn of talk during the interview, which can be a

sentence or a word (arguably the smallest grain size possible).

- **SS**: identifies semiotic systems, their affordances, and how they are combined for meaning-making. Every representation is considered an individual grain. This results in a small to intermediate grain size since this can be a whole drawing, but also a part of a drawing as it is constructed.

Of course, given different aims that lead to different analysis grain sizes and methodologies, the same data lead to vastly different results.

- **FWT**: A new misconception was suggested: The Moon looks larger at the equator than at the poles because the equator is closer to the Moon.
- **KiP**: Applying an existing p-prim (Ohm's p-prim) to a new context (astronomical distances) is suggested: The apparent changing size of the Moon is due to different locations on Earth (when one is on the equator and closer to the Moon it appears larger than when one is farther from the Moon in Sweden).
- **SS**: When the student is confident in how to describe the phenomena, a transductive chain is suggested that describes how the student distributes the meaning-making across multiple semiotic systems.

In FWT Mark's mental model stands firmly at the center of the researchers' attention. These authors tried to build a model as Mark might have seen it in his minds-eye (Fig. 5). The FWT researchers evaluate this model as relatively coherent and stable, since it appears over and over again in different questions. This is in line with what FWT would predict. However, the FWT researchers acknowledge, as stated above, that Mark is initially hesitant. He might thus be constructing the model during the interview. There is a lot of literature on the development of conceptual understanding and changes to mental models. However, the timescales in this literature are very different, years rather than minutes. Indeed, Vosniadou *et al.* [4] said that

the process of conceptual change is a gradual and continuous process that involves many interrelated pieces of knowledge and requires a long time to be achieved. (p. 26)

The rapidly evolving nature of Mark's conceptual understanding is thus unexpected from a strictly FWT point of view.

Rapid changes in conceptual understanding are, however, squarely within the ballpark of KiP. KiP research is interested in the small resources that might be used to construct an understanding. This construction can lead to a more or less coherent understanding (note that we do not use "mental model" here on purpose), but from the KiP perspective, the analytical focus rests on identifying and describing those resources and the process of organizing them and then reorganizing them over time. The mental

model of Mark was not the primary interest of the KiP researcher in this work. The “resources” referred to are, primarily, knowledge resources (e.g., p-prims).

However, there are in this interview also physical resources that Mark uses (both because it is required, cfr. the Suns and Moons on sticks, but also initiated by Mark, e.g., drawings). The way Mark interacts with these (semiotic) resources, and creates own (semiotic) resources (e.g., gesturing), as well as how he uses these resources to build an understanding and express his ideas, is the interest of the SS researchers.

A first notable point is thus that for all three models, a detailed analysis of the interview led to valuable insights that might be of interest to the respective research communities, despite their differences in interests, assumptions, aims, and methodologies.

While looking for adding to the literature based on each model alone was not the intention of this work, we showed that there is at least the possibility of gaining additional insights by analyzing (interview) data from different models. It can be expected that if yet other models are used even other insights might surface. This is not surprising, we do not think this is a revolutionary insight given the existing robust literature pertaining to each of these models. Instead, this small contribution is a (welcome) byproduct of the way the interview was selected out of three interviews that were conducted. This might not be true if we focused our attention to the other interviews.

B. The complementary nature of the three models

It is clear that all three models approach the mental imagery of Mark in a different way. It is also clear that none of them has full access. A second, and more important, key insight of this work is that the three models seem to be complementary in nature. Indeed, by using three different models to look at the same data we gained a more detailed and nuanced picture of Mark’s understanding of the astronomical phenomena in question.

We interpret the data, combining all three models, as follows, see Fig. 7. Mark got a question that he, prior to the interview, never gave dedicated thought. In Fig. 7, at time

t_0 p-prims, resources and prior mental structures existing in Marks’ mind are depicted as unconnected points, circles, and polygons. Mark is asked a first question and, at time t_1 , he hesitates but gives an initial answer, combining some p-prims, resources, and underlying mental structures. The connections are depicted as dashed lines, indicating that they are weak and tentative. However, as the interview progresses, these connections grow stronger (indicated by the solid lines in the mind on the right, at t_{end}) and incorporate new aspects, induced by the series of interview questions. The connections that he makes using various resources and p-prims are the area of interest of KiP (indicated in blue). Inversely, the final (semi-)stable, yet erroneous, structure that recurred throughout the interview is of interest for FWT (indicated in orange). Of course, both models take the whole interview into account (i.e., both colors are present from left to right). It is, however, not entirely clear what happens throughout the interview (from t_1 until t_{end}). This is indicated by the white, leaving room for and acknowledging the need for other approaches. The SS model (indicated in green) is one that does exactly that. Throughout the interview, Mark speaks, gestures, and draws. These actions of transducing his mental image to a physical mode and back force him to clarify his thinking. This strengthens the connections in his mind.

This picture is, of course, not complete., e.g., we do not know why the localized p-prim was triggered, perhaps because of the language barrier and an erroneous translation. We do not know how stable this mental model is. We do not know if and how it ties in to his understanding of other astronomical (and Earthly) phenomena (e.g., seasons, tides, day and night,...). Still, this interpretation is richer than any of the three models could achieve individually.

Finally, we note that, despite the differences, there are aspects that all three researcher-teams agree on. We agreed that Mark’s was the most interesting interview (though for different reasons), we agreed on which part of Mark’s interview was most interesting and we agree that Mark is initially hesitant but eventually builds a more or less coherent understanding.

C. Theoretical ideas that come from working together

So far, we discussed how the insights were relevant to three models’ individual literature, as well as the complementary nature of the models which allowed to add more detail to Mark’s case. A final contribution comes from the process of working together. Throughout the discussions, juxtaposing various combinations of the models, more nuanced insights and new questions arise.

1. KiP-SS

The SS emphasis on persistent and nonpersistent semiotic resources raises a productive question: how do KiP

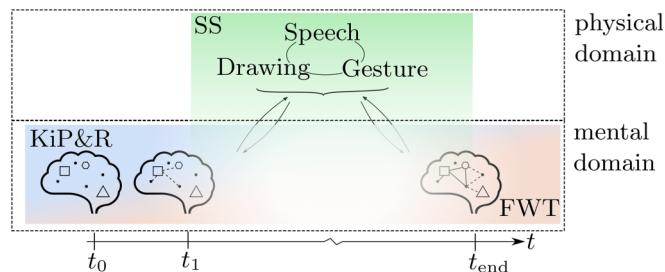


FIG. 7. The three models combine in a complementary way to describe how Mark’s mental imagery grew over time. The figure is described in detail in the text.

and SS each conceptualize the relationship between representational resources and knowledge? More specifically, might there be a systematic relationship between representational gestures and changes in a knowledge system? Although empirical KiP studies have incorporated gestures and representations into their analyses [87,88], explicitly integrating the concept of transduction could offer new analytical purchase on the mechanics of learning, illuminating not just what resources students draw on, but how meaning is remade as it moves between modes.

A related point of contact concerns verbal hesitation. Within a KiP analysis, hesitations are typically treated as evidence that a student is constructing an explanation in real time, drawing on and assembling available knowledge resources. From an SS perspective, the same hesitation could be read as a moment of attempted transduction, a pause in which meaning is being reshaped as it crosses from one semiotic system to another. These interpretations are not mutually exclusive; in fact, their convergence suggests that hesitation may mark a genuine cognitive-semiotic boundary.

Finally, while SS does not operate with an equivalent to the fine-grained “grain size” of phenomenological primitives (p-prims) in KiP, a productive relationship between the two levels can still be articulated. A given semiotic resource, whether a gesture, a diagram, or a verbal formulation, can be understood as a particular instantiation of some combination of p-prims. Crucially, however, no single semiotic mode can fully encode everything that such a combination carries, as some meaning is inevitably lost or transformed in the remaking. This is precisely what transduction describes. Seen this way, transduction can serve as the bridging concept that accounts for how p-prims become externalized in the physical environment, as a movement from internal knowledge structure to material semiotic form, while the need for multiple semiotic systems reflects the irreducible complexity of the underlying knowledge resources being expressed.

2. KiP-FWT

An existing body of work has already discussed many of the points of overlap and contention between the KiP and FWT model. The two “camps” discuss this in their own papers, see e.g. [7,8], but also outsiders refer to both as opposites in the coherence debate. Some have argued that the different stances are, not in a small part, due to methodological differences. Indeed, FWT studies typically have large N and often younger children, whereas small N and undergrads are more typical in KiP studies. That difference has been eliminated in this article since we use the exact same interview data. Still, the author teams came to different conclusions.

This, we think, is the result of a different focus and a different grain size of interest. The KiP analysis focused on

a smaller grain size as relevant to the context specificity of knowledge resources while the FWT focused on a longer timescale (still within the interview) to explain semi-stable conceptual understanding. These differences are, of course, nuanced as KiP recognizes that (in)consistencies can be important, and FWT also values the context, but the different emphasizes perpetuated into different results. As such, we think that different stances in the coherence debate might be, in part, a false dichotomy. We follow Scherr’s call for researchers to be open to both models and look for complementarity [24].

If students can both employ p-prims (and other resources) to construct answers to conceptual questions and build (semi-)coherent mental models, then this leads to interesting questions. What is the smallest internal structure that can be integrated into a mental model, and can p-prims/resources offer insights into this? When does a mental structure transition from temporary (question-induced) to more permanent? How much time (of active student engagement with learning materials/teachers) is needed to allow such a transition? Can such transitions be guided, toward a scientific understanding, by offering specific (external) resources or priming different p-prims?

3. FWT-SS

In FWT, gestures are typically analyzed as a window onto participants’ mental models, that is, they serve an observational function for the researcher. However, our analysis from SS model suggests that gestures also play a constitutive role: they are not merely expressions of preexisting understanding but are inherently part of the process through which Mark develops his conceptual understanding. We thus call for more collaborative efforts between these two traditions, possibly enriched by ideas from embodied cognition.

A second point of attention concerns the role of physical tools in the interview setting. While the interviews conducted by Vosniadou and Brewer [2,3] did not involve physical equipment, the interview analyzed here did, and we argue that the tools present had a substantive impact on Mark’s thinking and, consequently, on his responses. We concur with Schoultz *et al.* [31] that when equipment is used in interviews, its role must be carefully examined rather than treated as a neutral backdrop. Here, too, a social semiotic perspective proves valuable, given its explicit attention to how people marshal available resources to express meaning, a key element in any interview situation.

For both gestures and physical tools, but also for speech, Mark works through a process of transducing meaning in order to externalize his inner mental model. This is a process with limitations, no semiotic system can fully represent an idea and every transduction foregrounds some aspects while backgrounding others. We expect this to also hold for internalization—the process of transduction back

to inner mental imagery—which follows naturally from the very act of using and manipulating semiotic resources. Similar to transduction between several semiotic resources, internalization is prone to an overemphasis of certain aspects of the concept at hand and a loss of other aspects. As such, the very act of explaining, of trying to convey meaning, forms Mark's conceptual understanding. This aligns with a long-standing call for understanding conceptual change not merely as a cognitive process but as one that also can happen, and might be encouraged, through social interactions [35]. It is conceivable that egocentric or inner speech [89] follows the same pathways of transducing abstract mental models to (inner) speech back to mental imagery.

Note here that we switched from mental model (a term from FWT) to mental imagery (a term from SS). One could propose that these terms describe the same internal structure but foreground different aspects of it. FWT is interested in the representational content (what is it a model of? what other mental structures (ideas/models/...) is it connected to?). Meanwhile SS is interested in the characteristics of the semiotic nature of this structure (what affordances does internal meaning-making have, and what is gained or lost in transduction to and from external semiotic systems?). Whether this amounts to a genuine difference in constructs or merely a difference in analytical focus remains an open question, but one whose resolution would be an important step toward any future integration of these models.

4. FWT-KIP-SS

When considering the juxtaposition of the three models, there are new questions about the space between the models. One consideration is to extend the two more cognitive perspectives (FWT and KIP) into the space of embodied cognition. Although there is already some work in this space (e.g. [90]), there are new questions about how to extend KIP and FWT to more comprehensively take into account SS. In parallel one can consider how to extend SS to take into account cognition.

Another consideration is the role of the provided tools (e.g., Suns and Moons on a stick) and created tools (e.g., drawing). The created tools were the most interesting from the SS perspective as it involves meaning-making through the creation of those tools. Regarding the provided tools, there was also the paper disk that Mark chose at the beginning of the interview, which was of interest to the FWT. This disk was built into the interview protocol, had some FWT associated assumptions behind it, and thus was part of the FWT analysis, but less relevant to the KIP analysis. This all raises questions about instances when tools are and are not relevant to FWT and KIP analyses, and how to more carefully discuss their roles in the theory, its assumptions, and interview approaches.

X. LOOKING FORWARD

In this work, we investigated the (emerging) understanding of one topic by one student in one context. It is one of only very few examples where one set of data get interpreted through multiple models. We think this is just a first step of an evolution that is necessary in science education research, and its discipline based relatives (e.g., physics education research); an evolution that has both a theoretical and experimental component.

Theoretically, we see an assignment to reduce the plethora of models (frameworks, theories,...). Right now there are many models, 86 according to Ref. [1], all claiming to describe and explain conceptual change or understanding. This is too many, we think it is necessary and possible to reduce this number. While we do not think it is possible to reduce all of conceptual change to a single theory, reducing the number of models would clarify research both conceptually and methodologically. It would also reduce the variability in teacher training.

This article is highlighting the importance of careful descriptions about the models so that scholars can make informed research decisions. What do the models claim? In what context do the models make these claims? What are the underlying assumptions of the models? What learning phenomenon are the models describing? Subsequently, it is necessary to compare the models. In which aspect are models unique? What elements are shared over many models? Which models work better under which conditions? Finally, what are the common learning, teaching, and instructional implications that transcend multiple models? How can the multitude of models be combined into a limited number of models that are able to describe, explain, and predict conceptual change. Ideally, these combined models do so in a way that is both broad enough to allow for the large variations in contexts and contents, but also detailed enough to allow for experimentally verifiable predictions (through concrete hypotheses).

This theoretical work of juxtaposing three models in a cooperative collaboration has led to more refined questions and hypotheses that will move the field forward. Importantly, by using a qualitative collaboration, we are beginning to answer questions and generate new hypotheses that go beyond the prior attempts at a quantitative replication [15,17]. We see the value in both approaches given the complexity of educational researching in working to build pedagogical implications that can serve the field.

This theoretical work will have to be complemented by experimental work. For quantitative work progress has been made after the replication crisis [91] to structure collaboration. However, given the complexity of education, we think that there is a lot of foundational qualitative work needed that will allow us to build specific/stronger/refined hypotheses that cut across the different models

(in the future, these cross-cutting hypotheses have to be tested further, perhaps in a quantitative way). Indeed, in qualitative research, collaboration and replication are needed [92]. So, what might such a gold standard of scientific collaboration look like in qualitative work?

We referred to our approach as “cooperative collaboration,” differentiating it from “adversarial collaboration” [13]. In it different research groups (or representatives of a model) come together (i) to look at the same data, thereby eliminating slight variations in context that can be expected if similar data are used (e.g., data collected by each partner in their own context), and (ii) with the aim to look for complementarity. We think this kind of cooperative collaboration in qualitative research is necessary as a first step to get a grasp on the possibilities but also limitations of the different models, laying the ground works for eventual convergence and subsequent quantitative work.

Our research process for enacting this collaboration was iterative in nature. We first interpreted the data alone, then engaged in iterative rounds of group discussion among the author teams to go beyond individual analyses and bring results and questions together in a way that resulted in an insights that cut across the models. Once the article was drafted, having separate pieces read and commended on by authors from different models helped refine and crystallize the key points.

It is important to note in hindsight opportunities for improvement. First, although the interview procedure was agreed upon by all authors, it was initially developed by the FWT authors for another study (that sits within the FWT tradition). The emphasis on the scale model was build from an FWT assumption, and strongly incorporated into the FWT analysis, while being mostly overlooked in the KiP analysis. Although, it would likely be impossible to have neutral interview questions, recognizing the assumptions behind the questions is important in considering how the questions impact the analysis and results. Retrospectively, it would have been better to build the interview procedure from scratch to accommodate all three models. Second, it would have been beneficial to take the SS perspective into account when considering the role of gestures and tools in building the interview protocol. Third, we should not neglect the limitations inherent to

the used two-dimensional scale model. While practical for the research setup of this work, a “flat” scale model is hardly ideal to explain intrinsically three-dimensional phenomena such as Moon phases. One could critically argue that the use of such a scale model induces confusion between, e.g., a lunar eclipse and a full Moon phase. Although Mark did not seem to be led astray by this limitation, one should take it into account nonetheless.

Finally, we consider the pedagogical/didactic implications. First, from the perspective of all three models, we recognize the importance of sufficient time to allow these reasoning processes to occur. We also recognize the value in tools, gestures, and other artifacts that impact this reasoning. From the perspective of astronomy didactics, these results suggest the potential to have students reflect on their position on Earth and whether that might inadvertently impact their reasoning about astronomical phenomena. These results also contribute to the widely known concern around students struggling with astronomical distances.

XI. OVERARCHING CONCLUSION

In this work, we analyzed an interview with one student from three models: framework theory, knowledge-in-pieces/resources and social semiotics. The analysis in each model led to a contribution to their respective literatures: a new alternative conception was uncovered (FWT), there is evidence in favor of localized p-prims (KiP) and evidence of a transductive chain (SS). However, by directly comparing how the same data are analyzed in three different ways the complementary nature of these perspectives became evident. Moreover, the shared analysis led to deeper insights, allowing each model to progress further. We think this is the real added value of this work and argue for more direct comparisons of models using shared data. We argue that this is a necessary step to progress the field.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

-
- [1] P. Potvin, L. Nenciovici, G. Malenfant-Robichaud, F. Thibault, O. Sy, M. A. Mahhou, A. Bernard, G. Allaire-Duquette, J. Blanchette Sarrasin, L.-M. Brault Foisy *et al.*, Models of conceptual change in science learning: Establishing an exhaustive, *Stud. Sci. Educ.* **56**, 157 (2020).
 - [2] S. Vosniadou and W.F. Brewer, Mental models of the earth: A study of conceptual change in childhood, *Cognit. Psychol.* **24**, 535 (1992).
 - [3] S. Vosniadou, Capturing and modeling the process of conceptual change, *Learn. Instr.* **4**, 45 (1994).

- [4] S. Vosniadou, in *International Handbook of Research on Conceptual Change* (Routledge, New York (NY), 2008), pp. 31–62.
- [5] A. A. diSessa, Toward an epistemology of physics, *Cognit. Instr.* **10**, 105 (1993).
- [6] T. G. Amin and O. Levriani, *Converging Perspectives on Conceptual Change: Mapping an Emerging Paradigm in the Learning Sciences* (Routledge, New York (NY), 2018).
- [7] A. A. diSessa *et al.*, in *The Cambridge Handbook of the Learning Sciences* (Cambridge University Press, NY, 2014), pp. 88–108.
- [8] S. Vosniadou, The development of students' understanding of science, *Front. Educ. (Lausanne)* **4**, 32 (2019).
- [9] T. G. Amin, *Learning Scientific Concepts: A Unified Theory of Conceptual Change* (Taylor & Francis, London, 2025).
- [10] P. Johnson-Laird, *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness* (Harvard University Press, Cambridge (Massachusetts), 1983).
- [11] J. Airey and C. Linder, *Multiple Representations in Physics Education* (Springer, New York, 2017), pp. 95–122.
- [12] M. Kersting, T. G. Amin, E. Euler, B. Gregorcic, J. Haglund, L. K. Hardahl, and R. Steier, What is the role of the body in science education? A conversation between traditions, *Sci. Educ.* **33**, 1171 (2024).
- [13] B. Mellers, R. Hertwig, and D. Kahneman, Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration, *Psychol. Sci.* **12**, 269 (2001).
- [14] Oxford University Press, Coopetition, n. in Oxford English dictionary (n.d.). retrieved February 24, 2026, from [10.1093/oed/4424191723](https://www.oed.com/entry/10.1093/oed/4424191723).
- [15] A. A. diSessa, N. M. Gillespie, and J. B. Esterly, Coherence versus fragmentation in the development of the concept of force, *Cogn. Sci.* **28**, 843 (2004).
- [16] C. Ioannides and S. Vosniadou, The changing meanings of force, *Cognit. Sci. Q.* **2**, 5 (2002).
- [17] D. B. Clark, C. M. D'Angelo, and S. P. Schleigh, Comparison of students' knowledge structure coherence and understanding of force in the Philippines, Turkey, China, Mexico, and the United States, *J. Learn. Sci.* **20**, 207 (2011).
- [18] S. Vosniadou and I. Skopeliti, Conceptual change from the framework theory side of the fence, *Sci. Educ.* **23**, 1427 (2014).
- [19] D. T. Brookes and E. Etkina, in *Conceptual Metaphor and Embodied Cognition in Science Learning* (Routledge, New York (NY), 2018), pp. 27–47.
- [20] A. A. diSessa, Conceptual change in a microcosm: Comparative learning analysis of a learning event, *Hum. Dev.* **60**, 1 (2017).
- [21] M. T. H. Chi, J. D. Slotta, and N. De Leeuw, From things to processes: A theory of conceptual change for learning science concepts, *Learn. Instr.* **4**, 27 (1994).
- [22] A. Elby, What students' learning of representations tells us about constructivism, *J. Math. Behav.* **19**, 481 (2000).
- [23] E. Blown and T. G. K. Bryce, Conceptual coherence revealed in multi-modal representations of astronomy knowledge, *Int. J. Sci. Educ.* **32**, 31 (2010).
- [24] R. E. Scherr, Modeling student thinking: An example from special relativity, *Am. J. Phys.* **75**, 272 (2007).
- [25] D. B. Clark, Longitudinal conceptual change in students' understanding of thermal equilibrium: An examination of the process of conceptual restructuring, *Cognit. Instr.* **24**, 467 (2006).
- [26] J.-M. G. Rodriguez and M. H. Towns, Analysis of biochemistry students' graphical reasoning using misconceptions constructivism and fine-grained constructivism: Why assumptions about the nature and structure of knowledge matter for research and teaching, *Chem. Educ. Res. Pract.* **22**, 1020 (2021).
- [27] B. Farlow, M. Vega, M. E. Loverude, and W. M. Christensen, Mapping activation of resources among upper division physics students in non-Cartesian coordinate systems: A case study, *Phys. Rev. Phys. Educ. Res.* **15**, 020125 (2019).
- [28] M. Vega, W. Christensen, B. Farlow, G. Passante, and M. Loverude, Student understanding of unit vectors and coordinate systems beyond Cartesian coordinates in upper division physics courses, *J. Undergrad. Res. Nat. Sci. Math.* **19**, 194 (2017).
- [29] E. F. Redish, Oersted lecture 2013: How should we think about how our students think? (2014).
- [30] C. R. Samuelsson, Resourceful Students: Engaging Students in Active and Systematic Investigations in Laboratories Involving Thermal Phenomena, Ph.D. thesis, Acta Universitatis Upsaliensis, 2023.
- [31] J. Schoultz, R. Säljö, and J. Wyndhamn, Heavenly talk: Discourse, artifacts, and children's understanding of elementary astronomy, *Hum. Dev.* **44**, 103 (2001).
- [32] J. Bernhard, What matters for students' learning in the laboratory? Do not neglect the role of experimental equipment!, *Instr. Sci.* **46**, 819 (2018).
- [33] O. Halldén, L. Haglund, and H. Strömdahl, Conceptions and contexts: On the interpretation of interview and observational data, *Educ. Psychol.* **42**, 25 (2007).
- [34] L. Mason, Introduction: Bridging the cognitive and socio-cultural approaches in research on conceptual change: Is it feasible?, *Educ. Psychol.* **42**, 1 (2007).
- [35] P. R. Pintrich, R. W. Marx, and R. A. Boyle, Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change, *Rev. Educ. Res.* **63**, 167 (1993).
- [36] G. M. Sinatra, The "warming trend" in conceptual change research: The legacy of Paul R. Pintrich, *Educ. Psychol.* **40**, 107 (2005).
- [37] K. Svensson and E. Campos, Comparison of two semiotic perspectives: How do students use representations in physics?, *Phys. Rev. Phys. Educ. Res.* **18**, 020120 (2022).
- [38] I. Seidman, *Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences* (Teachers College Press, New York (NY), 2006).
- [39] J. A. Hatch, *Doing Qualitative Research in Education Settings* (SUNY Press-State University of New York Press, New York (NY), 2023).
- [40] K. Willem, D. C. Mieke, V. W. Hans, V. Wim, and J. Sermeus, The effect of scale comprehension on student reasoning about astronomical phenomena. I. The apparent sizes of Moon and Sun and solar eclipses, *Phys. Rev. Phys. Educ. Res.* **22**, 020106 (2026).

- [41] K. Willem, D.C. Mieke, V.W. Hans, V. Wim, and J. Sermeus, Effect of scale comprehension on student reasoning about astronomical phenomena. II. Moon phases, *Phys. Rev. Phys. Educ. Res.* **22**, 020107 (2026).
- [42] W. Keppens, M. De Cock, H. Van Winckel, W. Van Dooren, and J. Sermeus, Exploring student estimates of astronomical scales: Impact of question formulation and visualization, *Phys. Rev. Phys. Educ. Res.* **21**, 010159 (2025).
- [43] C. Linder, A challenge to conceptual change, *Sci. Educ.* **77**, 293 (1993).
- [44] A. Lelliott and M. Rollnick, Big Ideas: A review of astronomy education research 1974–2008, *Int. J. Sci. Educ.* **32**, 1771 (2010).
- [45] S. Salimpour, M. Fitzgerald, and U. Eriksson, A descriptive overview of English-language publications in the field of Astronomy Education Research, 1898 to 2022, *Astron. Educ. J.* **4**, 140aer (2024).
- [46] L. Shore and R. Kilburn, The effect of astronomy teaching experience on the astronomy interest and conceptions of elementary school teachers, in *Proceedings of the 3rd International Seminar on Misconceptions and Educational Strategies in Science and Mathematics* (Boston University, Massachusetts, 1993), p. 1.
- [47] P. Sadler, The Initial Knowledge State of High School Astronomy Students, Ph.D. thesis, Harvard University, Cambridge, MA, 1992.
- [48] B.W. Miller and W.F. Brewer, Misconceptions of astronomical distances, *Int. J. Sci. Educ.* **32**, 1549 (2010).
- [49] U. Kanli, A study on identifying the misconceptions of pre-service and in-service teachers about basic astronomy concepts, *Eurasia J. Math. Sci. Technol. Educ.* **10**, 471 (2014).
- [50] M. LoPresto and S. Murrell, An astronomical misconceptions survey, *J. Coll. Sci. Teach.* **40**, 14 (2011).
- [51] M. Zeilik and V.J. Morris, An examination of misconceptions in an astronomy course for science, mathematics, and engineering majors, *Astron. Educ. Rev.* **2**, 101 (2003).
- [52] W. Bisard, R. Aron, M. Francek, and B. Nelson, Assessing selected physical science and earth science misconceptions of middle school through university preService teachers, *J. Coll. Sci. Teach.* **24**, 38 (1994).
- [53] S. Vosniadou, in *Converging Perspectives on Conceptual Change. Mapping an Emerging Paradigm in the Learning Sciences*, edited by T. Amin and O. Levrini (Routledge, London and New York, 2018), Chap. 2, pp. 17–25.
- [54] V. Braun and V. Clarke, Using thematic analysis in psychology, *Qual. Res. Psychol.* **3**, 77 (2006).
- [55] I. Starakis and K. Halkia, Addressing k-5 students' and pre-service elementary teachers' conceptions of seasonal change, *Phys. Educ.* **49**, 231 (2013).
- [56] D. Hammer, Student resources for learning introductory physics, *Am. J. Phys.* **68**, S52 (2000).
- [57] A.A. diSessa and B.L. Sherin, What changes in conceptual change?, *Int. J. Sci. Educ.* **20**, 1155 (1998).
- [58] K. Jelacic, M. Planinic, and G. Planinsic, Analyzing high school students' reasoning about electromagnetic induction, *Phys. Rev. Phys. Educ. Res.* **13**, 010112 (2017).
- [59] S. Kapon and A.A. diSessa, Reasoning through instructional analogies, *Cognit. Instr.* **30**, 261 (2012).
- [60] H.C. Sabo, L.M. Goodhew, and A.D. Robertson, University student conceptual resources for understanding energy, *Phys. Rev. Phys. Educ. Res.* **12**, 010126 (2016).
- [61] D.E. Young and D.C. Meredith, Using the resources framework to design, assess, and refine interventions on pressure in fluids, *Phys. Rev. Phys. Educ. Res.* **13**, 010125 (2017).
- [62] K.S. Taber and A. García-Franco, Learning processes in chemistry: Drawing upon cognitive resources to learn about the particulate structure of matter, *J. Learn. Sci.* **19**, 99 (2010).
- [63] J.H. Doherty, J.A. Cerchiara, E.E. Scott, L.N. Jescovitch, J.L. McFarland, K.C. Haudek, and M.P. Wenderoth, Oaks to arteries: The physiology core concept of flow down gradients supports transfer of student reasoning, *Adv. Physiol. Educ.* **47**, 282 (2023).
- [64] J.-P. Burde and T. Wilhelm, Teaching electric circuits with a focus on potential differences, *Phys. Rev. Phys. Educ. Res.* **16**, 020153 (2020).
- [65] E.J. Blown and T.G.K. Bryce, Knowledge restructuring in the development of children's cosmologies, *Int. J. Sci. Educ.* **28**, 1411 (2006).
- [66] O. Parnafes, Developing explanations and developing understanding: Students explain the phases of the moon using visual representations, *Cognit. Instr.* **30**, 359 (2012).
- [67] B.L. Sherin, M. Krakowski, and V.R. Lee, Some assembly required: How scientific explanations are constructed during clinical interviews, *J. Res. Sci. Teach.* **49**, 166 (2012).
- [68] E.C. Sayre and M.C. Wittmann, Plasticity of intermediate mechanics students' coordinate system choice, *Phys. Rev. ST Phys. Educ. Res.* **4**, 020105 (2008).
- [69] A. DiSessa, B. Sherin, and M. Levin, *Knowledge and Interaction: A Synthetic Agenda for the Learning Sciences*, edited by A. diSessa, M. Levin, and N. Brown (Routledge, NY, 2016), Vol. 30.
- [70] L.A. Barth-Cohen, H. Swanson, and J. Arnell, Methods of research design and analysis for identifying knowledge resources, *Phys. Rev. Phys. Educ. Res.* **19**, 020119 (2023).
- [71] L. Barth-Cohen, Threads of local continuity between centralized and decentralized causality: Transitional explanations for the behavior of a complex system, *Instr. Sci.* **46**, 681 (2018).
- [72] D. Hammer, More than misconceptions: Multiple perspectives on student knowledge and reasoning, and an appropriate role for education research, *Am. J. Phys.* **64**, 1316 (1996).
- [73] M. Bang and D. Medin, Cultural processes in science education: Supporting the navigation of multiple epistemologies, *Sci. Educ.* **94**, 1008 (2010).
- [74] T.M. Philip, An "ideology in pieces" approach to studying change in teachers' sensemaking about race, racism, and racial justice, *Cognit. Instr.* **29**, 297 (2011).
- [75] T. Fredlund, J. Airey, and C. Linder, Enhancing the possibilities for learning: Variation of disciplinary-relevant aspects in physics representations, *Eur. J. Phys.* **36**, 055001 (2015).

- [76] T. S. Volkwyn, Learning physics through transduction: A social semiotic approach, Ph.D. thesis, Acta Universitatis Upsaliensis, 2020.
- [77] U. Eriksson, Reading the sky: From starspots to spotting stars, Ph.D. thesis, Acta Universitatis Upsaliensis, 2014.
- [78] J. Bezemer and G. Kress, Writing in multimodal texts: A social semiotic account of designs for learning, *Written Commun.* **25**, 166 (2008).
- [79] K. Svensson and U. Eriksson, Concept of a transductive link, *Phys. Rev. Phys. Educ. Res.* **16**, 026101 (2020).
- [80] T. Fredlund, J. Airey, and C. Linder, Exploring the role of physics representations: An illustrative example from students sharing knowledge about refraction, *Eur. J. Phys.* **33**, 657 (2012).
- [81] B. Gregoric, G. Planinsic, and E. Etkina, Doing science by waving hands: Talk, symbiotic gesture, and interaction with digital content as resources in student inquiry, *Phys. Rev. Phys. Educ. Res.* **13**, 020104 (2017).
- [82] C. Goodwin, *Discourse, the Body, and Identity* (Springer, New York, 2003), pp. 19–42.
- [83] ELAN (Version 7.1) [Computer software], Nijmegen: Max Planck institute for psycholinguistics, 2026, The Language Archive. Retrieved from <https://archive.mpi.nl/tla/elan>.
- [84] M. W. Alibali and M. J. Nathan, Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures, *J. Learn. Sci.* **21**, 247 (2012).
- [85] R. E. Scherr, Gesture analysis for physics education researchers, *Phys. Rev. ST Phys. Educ. Res.* **4**, 010101 (2008).
- [86] J. Sweller, *Psychology of Learning and Motivation* (Elsevier, New York, 2011), Vol. 55, pp. 37–76.
- [87] O. Parnafes, What does “fast” mean? Understanding the physical world through computational representations, *J. Learn. Sci.* **16**, 415 (2007).
- [88] L. A. BARTH-COHEN and M. C. Wittmann, Aligning coordination class theory with a new context: Applying a theory of individual learning to group learning, *Sci. Educ.* **101**, 333 (2017).
- [89] J. Piaget, Commentary on Vygotsky's criticisms of language and thought of the child and Judgement and reasoning in the child, *New Ideas Psychol.* **18**, 241 (2000).
- [90] S. Kapon, Unpacking sensemaking, *Sci. Educ.* **101**, 165 (2017).
- [91] M. Korbacher, F. Azevedo, C. R. Pennington, H. Hartmann, M. Pownall, K. Schmidt, M. Elsherif, N. Breznau, O. Robertson, T. Kalandadze *et al.*, The replication crisis has led to positive structural, procedural, and community changes, *Commun. Psychol.* **1**, 3 (2023).
- [92] M. C. Makel, M. S. Meyer, M. A. Simonsen, A. M. Roberts, and J. A. Plucker, Replication is relevant to qualitative research, *Educ. Res. Eval.* **27**, 215 (2022).