Probing students' understanding of some conceptual themes in general relativity

Atanu Bandyopadhyay[*] and Arvind Kumar[†]

Homi Bhabha Centre for Science Education, Tata Institute of Fundamental Research, Mankhurd, Mumbai 400088, India
(Received 18 November 2009; published 28 July 2010)

This work is an attempt to see how physics undergraduates view the basic ideas of general relativity when they are exposed to the topic in a standard introductory course. Since the subject is conceptually and technically difficult, we adopted a "case studies" approach, focusing in depth on about six students who had just finished a one semester course on special relativity. The methodology of investigation involved a combination of text comprehension questionnaire and detailed clinical interviews. The aim was not to investigate the technical proficiency of the students, but to probe in detail the nuances of their conceptions of several basic points of the subject. Analysis of their responses reveals a large number of "alternative conceptions" of students in the domain. The study should be useful to physics education researchers as well as to teachers of introductory general relativity at about the senior undergraduate level.

I. INTRODUCTION

This work is part of a program of cognitive studies on relativity being pursued at our institute in Mumbai. A study on nearly all aspects of Galilean relativity was carried out more than a decade ago (see [1–3]). More recently, we examined how students view the different but equivalent "meanings" (i.e., formulations) of the relativity principle in Newtonian mechanics [4]. In the present study, we probe in detail how physics undergraduates view the basic ideas of general relativity. In a parallel investigation (to be reported elsewhere), the focus is on students' understanding of the kinematics of special relativity (SR).

The general framework of alternative conceptions' research that informs this study has been extensively employed in the literature. This approach, rooted in the constructivist paradigm of learning, posits that prior knowledge and conceptions of students interfere with and affect their learning of new domains. Alternative conceptions are (loosely) organized knowledge structures (schemata) that often arise from the spontaneous ideas that students acquire in their normal cognitive development through interaction with the environment. They may also arise from (or get modified by) instruction or from prior knowledge of antecedent topics. Several of students' alternative conceptions are found to be universal and robust i.e., they are resistant to instruction and may even coexist with standard conceptions learnt in a classroom. Common examples are the notions of force, heat, and temperature and ideas of vision. Research within this framework generally, if not always, deals with domains that do not demand highly technical or mathematical prerequisites and are amenable to qualitative analysis. (For access to this kind of research, see [5] and the references therein).

General relativity (GR) is known to be a conceptually and technically intricate subject. The conceptual and the technical are often inseparable in advanced domains of physics. Yet Einstein's relativity is one domain where this separation is meaningfully possible, to a great extent in SR and lesser but significant extent in GR. Our entire program is premised on this possibility. The emphasis on the qualitative aspects, however, does not mean that we neglect the essential technical aspects of the subject. The idea is to teach the subject in all its technical detail appropriate at the senior undergraduate level, but concentrate on probing students' grasp of the conceptual content of the domain. Earlier pedagogic work on GR has mainly dealt with suggestions on how to make the learning of the mathematical and physical content of GR more accessible to students; see, for example, Refs. [6,7]. There are, of course, several papers that deal with pedagogic clarifications, alternative derivations, etc. of particular topics in GR. See, for example, Refs. [8,9]. However, a detailed diagnostic study of the kind reported here does not seem to have been carried out.

We should add that there is considerable debate in the literature regarding how best to characterize and deal with students' patterns of errors in physics (see [10,11]). The different theoretical frameworks view the process of knowledge acquisition in subtly different ways. Further, ontological presuppositions and epistemological beliefs are also known to affect (and be affected by) the learning process. We do not get into these issues here, since our study focuses only on diagnosing students' conceptions and not on conceptual change in view of the complexity of the domain. In this study, for consistency of usage, we call all such student conceptions that are not congruent with standard conceptions of physics as "alternative conceptions."

II. METHODOLOGY

The methodology of the study had to be customized for the specific domain at hand. The details are as follows.

A. Design

The design adopted a case studies approach to enable us to carry out an in-depth diagnosis of students' conceptions. The standard method of designing diagnostic problem tasks and studying students' responses to them did not seem ap-

---

*atanu@hbcse.tifr.res.in
†arvindk@hbcse.tifr.res.in

propriate for this design. We needed to probe students' ideas as they began to interact with the subject and its radically new notions. The method that suggested itself was to use text comprehension questionnaire as a tool of the study at the beginning of the course. In the standard teaching course on GR that followed, special questionnaires on some key notions were administered. The study ended with detailed interviews of all the students at the end of the course, each recorded interview lasting for two hours or so in which we essentially revisited the same text and asked students similar (but not identical) questions as in the text comprehension questionnaire.

B. Text comprehension

A little reflection suggested that the ideal text for our cognitive study of relativity was among the best known books of the world, written by none other than the originator of the subject himself (Einstein A., *Relativity: The Special and General Theory*) [12]. That this book is not a 'popular' exposition of relativity (in the sense of compromising the conceptual rigor of the subject) has been clearly stated by Einstein in the preface to the book:

"The present book is intended, as far as possible, to give an *exact* insight into the Theory of Relativity to those readers who…… are not conversant with the mathematical apparatus of theoretical physics" [our emphasis]

The book is eminently suitable for reading sessions and text comprehension questionnaire. In about a hundred pages it covers SR (articles I to XVII) and GR (articles XVIII to XXIX). Accordingly, the first step in our methodology was as follows: the instructor would read out an article of the book slowly, without explaining its "physics," but occasionally clarifying the meaning of a difficult English word or phrase. Students were then asked to give written responses to a few questions (usually three or four) based on the article. The questions were phrased in essentially two ways, requiring the student to either interpret some quotes of the text or clarify some key idea of the article. An example of the latter is: "what is the alternative interpretation of the observer in the railway carriage experiencing a jerk forward as a result of the application of the brakes?" (see Sec. III D). The questions did *not* involve going beyond the text or applying the ideas of the article to new situations. The full questionnaire consisted of 52 such items. For the first eight conceptual themes of Sec. III, these are given following the text summary under each theme.

C. Student sample

The sample for the study on GR consisted essentially of six undergraduate students (all boys, labeled here S1 to S6) who had finished high school two years ago. One of these (S6) did not write the text comprehension questionnaire but participated in the rest of the course including the end-of-course interviews. All the students had gone through a course on SR based on Robert Resnick's well-known book on the subject [13]. The competency levels of the students in general physics and SR were not tested at the beginning of the GR course, but this can be gauged from the following: (a)

they were among about twenty students admitted to the institute on the basis of a nationwide entrance test that had items of varying difficulty level, roughly comparable to the problems of Halliday, Resnick, and Walker's book [14], (b) out of the six students, three had got "A" grade, two "B" grade and one "C" grade in the preceding full-semester course on SR mentioned above, and (c) all of them had successfully completed other concurrent undergraduate physics courses in the last two years. The high level of motivation of the students was evident from the fact that the course on GR was 'on demand' and students cheerfully volunteered to forgo their long vacation for learning this new subject.

D. Teaching course on GR

Since general relativity was not part of the syllabus, a special three-month long vacation course was offered as an optional course. The course was short but intensive, entailing about seventy hours in all. The course began with the text comprehension questionnaire on the GR part of Einstein's text as explained earlier. It took two long sessions, each about four to five hours, on two consecutive days to administer the questionnaire on the twelve articles on general relativity. This was followed by a standard teaching course based on Ray D′ Inverno's introductory book on relativity [15]. The course covered tensor calculus and basic general relativity parts of the book, including the classic tests of GR and introductory cosmology. This ambitious coverage was possible mainly because students had essentially no other academic involvement during the vacation. The lectures were given in a relaxed and interactive mode throughout (each teaching session lasting about 4 h with a short break): they mainly involved careful working out of the mathematical steps of the book and explicating their physical interpretation, where relevant. Homework assignments generally consisted of three to four exercises aimed at filling in the gaps in the text or doing some of the relatively simple end-of-chapter computational problems. The conceptual issues highlighted in Einstein's text were not explicitly wrestled with either in the course or in the homework assignments. However, students did engage with qualitative aspects in three specially designed questionnaires on some key notions of GR: (a) the principle of equivalence, (b) the principle of general covariance and (c) space-time curvature, administered at appropriate times during the course. Although these questionnaires were not overly mathematical (as traditional tests on GR would be), they were far more technical and detailed than the broad conceptual themes of Einstein's text. In the end-of-course interviews, we returned to Einstein's text and probed how students, having had a technical exposure to the topic, engaged with the same conceptual issues they did in the beginning of the course. This paper is based on the text comprehension questionnaire and the end-of-course interviews. The analysis of students' responses to the technical questionnaires mentioned above will be reported elsewhere.

III. ANALYSIS OF STUDENTS' RESPONSES

We now analyze students' ideas on the basics of general relativity as evidenced by their written responses to the text

comprehension questionnaire and oral responses in the end-of-course interviews. For a meaningful presentation, we cluster the large number of items in the questionnaire into about nine key themes of general relativity described in Einstein's text. For the first eight themes (**A** to **H**), we begin by stating concisely the author's intended meaning (as far as we understand it), then state what the students were asked concerning it and finally describe their written and oral responses. The last theme (**I**) is rather technical and is treated very briefly. For a case studies approach, we should ideally narrate each student's responses separately. However, for compactness and coherence of presentation, we put these together in the form of our general impressions of their conceptions, with different nuances for different cases. Excerpts from the written responses and interviews are quoted where necessary [16]. Each theme ends with a "Commentary" in which students' conceptions on the theme are briefly summarized. Additionally, in a separate paragraph, we give some suggestions which we feel could be useful to the teachers of the subject. These pedagogic suggestions are not necessarily supported by our data.

The analysis is not focused on looking for conceptual change (which is a difficult undertaking even in simpler domains) but more on identifying students' alternative conceptions. These usually survive technical exposure to the subject. This is why we have juxtaposed the text comprehension questionnaire data with the interviews data to look for students' notions that stay through and after the standard course on the subject. In short we are aiming at a nuanced understanding of what it is that students understand reasonably well and what it is that is problematic for them, in regard to the basic notions of general relativity. In the last section we make some general observations on the possible sources of students' conceptions.

A. Relativity of uniform motion

Einstein begins his exposition of general relativity by recalling the special principle of relativity, i.e., relativity of uniform motion. Taking the familiar carriage-embankment example, he says that either of the two could be taken as a reference body and motion referred to it but this self evident assertion "must not be confused with the more comprehensive statement called the principle of relativity." The latter means that the general laws of nature (e.g., the laws of mechanics or the law of propagation of light in vacuum) have the same form in both the reference bodies. "Unlike the first, this latter statement need not of necessity hold a priori"—only experience can decide if it is correct or incorrect.

Students were asked to explain the underlined quotes above.

1. Students' ideas

Three students (S1, S4 and S5) quite clearly understood the point that the principle of relativity was more than the self-evident assertion above. S5 displayed his clear understanding while further stating that while "kinematic relativity" (his term for the self-evident assertion above) is true even for frames in relative nonuniform motion, the principle

of relativity is not true for them. The interviews confirmed these impressions.

S3 had not understood the phrase *a priori* in the script. When explained in the interview, he said after some struggle that the principle of relativity needed to be tested but it was *somewhat obvious*. He went on to say, however, that it was certainly not obvious for velocity-dependent forces (such as magnetic forces) and needed to be tested. It seemed that by "obvious," S3 meant what followed from daily experience in the realm of mechanics.

S6 seemed to think that the principle was obvious for translational motion but not so for more complicated motion (perhaps he had rotational motion in mind). Clearly, he did not realize that it was not obvious even for uniform translational motion.

"I think he is trying to say that when you see the embankment moving with respect to the carriage, or the carriage moving with respect to the embankment, the physical laws are obviously similar [in the two frames]. But when you talk of other frames with more complicate motion, the laws may not be similar" [S6]

S2 did not get what the self evident assertion was, though he did understand that the principle of relativity required the same form of laws for the two frames and that this fact needed to be tested experimentally. The interview confirmed this impression.

2. Commentary

Though all students could readily recite the principle of special relativity (laws have the same form in different inertial frames), only three of them clearly grasped the important point that this principle is not *a priori* true and requires experimental verification. This point is true even for Galilean relativity and students' analogous notions have been discussed in detail in our earlier work [17].

Instruction needs to emphasize that the relativity of uniform translational motion does not follow directly from the primitive notions of motion and frames, that it is much more than the self evident fact of being able to refer motion (i.e., describe phenomena) with respect to any body of reference, and that it is an experimentally verified fact of nature.

B. Nonuniformly moving frames of reference and the general principle of relativity

Now, Einstein says, the temptation to generalize the special principle of relativity to the general principle of relativity (the laws of nature hold for all bodies of reference, whatever be their state of motion) is natural. But there is a problem. For example, as long as the carriage is moving uniformly, "the occupant of the carriage is not sensible of its motion" (i.e., the same laws hold as for the embankment and he can just as well consider himself at rest and the embankment moving). But when the carriage moves nonuniformly, as when it is retarded by an application of brakes, mechanical behavior of bodies is different from that for the uniformly moving case. Galilean law, i.e., the First Law of motion no longer holds. "Because of this we feel compelled at the present juncture to grant a kind of absolute physical reality to

nonuniform motion, in opposition to the general principle of relativity."

Students were asked to state what Einstein meant by the general principle of relativity and explain the underlined quote above.

1. Students' ideas

Three students (S1, S3, and S4) stated the general principle of relativity quite correctly. "All frames of nature are equivalent for the description of natural laws irrespective of their states of motion." (S1, S3). Note the use of the word "laws" instead of "phenomena." Only S4 spoke of the form invariance of laws: "The same form of physical laws is applicable for inertial and noninertial frames." (S4).

Student S5 showed an unexpected idea: "By general principle of relativity Einstein means that only inertial frames should not be given privilege over others in the description of the laws of nature. Natural phenomena can be described using any reference frame, not necessarily an inertial frame." (S5). The student clearly did not realize that the last line was self-evident and did not need the general principle of relativity.

The interview of this student brought out the same point even more glaringly.

I: "What is the meaning of the general principle of relativity?"

S5: "It means every observer is capable of discovering the laws of physics staying in his own frame. It doesn't matter if he is in an inertial frame or a noninertial frame."

I: "Isn't that self-evident? Everybody can, of course, get the laws of physics relative to his frame."

S5: "No, that is not self-evident. For example, in Newtonian Mechanics or special relativity, there are only a few classes of observers who are capable of doing this thing, discovering the laws of physics."

This student knew clearly (elsewhere in the interview) that laws are different for inertial and noninertial frames in Newtonian mechanics. He showed good technical competence otherwise. Yet after a full course, he seemed to have his own epistemological stance regarding "discovering laws of physics." Variants of this stance were apparent in other cases too. S6 stated that GR meant there is no privileged frame of reference. But he put on a rider: the laws may be the same in all the frames, but for some (inertial frames) it may be easy to obtain the laws, for others like rotating or accelerating frames it may not be so easy.

As for the second question (explaining the underlined quote above), all students except S2, understood that the laws were different for nonuniformly moving frames from those in inertial frames, and that this was in conflict with the principle of general relativity. But only two of them (S1 and S5) stated that this implied 'absolute reality' to nonuniform motion. S5 stated it best:

"The acceleration of a frame can be found out by experiments carried out in that frame and thus an 'absolute' sense of its [nonuniform] motion can be granted to it." (S5)

2. Commentary

The distinction between the "description of phenomena" and the "laws of nature" so crucial to a proper understanding of relativity is often blurred among students, some of whom entertain notions of the "ease" (or the lack of it) in discovering the laws of physics in different frames. They readily understand that the laws are different in noninertial frames, but it is doubtful if they appreciate how fundamental this difficulty was for Einstein—namely, to reconcile "absolute physical reality of nonuniform motion" with his general principle of relativity, and how the Principle of Equivalence (see below) resolved the problem of absoluteness for him.

Instruction must clearly distinguish between the "description of phenomena" (which is nothing but measurement of the concerned physical quantities) and the "laws of nature" (general relations between the measurements), and help students discard anthropomorphic notions of "observer," "difficulty in discovering laws," etc. An observer in any reference body (frame) can, through measurements, arrive at the laws of nature relative to his frame. The laws may, of course, have more complicated forms (e.g., pseudoforces appear in noninertial frames) than the simple forms they have in inertial frames.

C. Equality of inertial and gravitational masses

Einstein introduces the notion of field for electromagnetism and gravity. He then points out the remarkable property of the gravitational field (not shared by electric or magnetic fields) namely, that "bodies which are moving under the sole influence of a gravitational field receive an acceleration, which does not in the least depend either on the material or the physical state of the body." From this experimental fact he concludes that the inertial mass (m_I) appearing in the law of motion $\vec{F} = m_I \vec{a}$ is proportional to (equal to, by choice) the gravitational mass (m_G) that appears in the gravitational law $\vec{F} = m_G \vec{g}$, where $\vec{g}$ is the intensity of the gravitational field or simply the field. A satisfactory understanding of this equality is possible, he says, if we recognize that *the same quality of a body manifests itself according to circumstances as "inertia" or "weight."*

Students were asked: (a) what was remarkable about the gravitational field in contrast to electric and magnetic fields, (b) whether the equality $m_I = m_G$ was *a priori* obvious, and (c) to explain the underlined statement above.

1. Students' ideas

(a) Property of gravitational field

Two students (S1 and S3) repeated essentially what was written in the text without explaining the contrast between a gravitational field and an electric or a magnetic field. S4 and S5 stated it more clearly,

"Acceleration produced by a gravitational field on a piece of wood or iron is the same. But a magnetic field attracts iron not wood." (S4).

"All bodies acquire the same acceleration in a gravitational field—in contrast electric and magnetic fields act differently on different bodies depending on charge, etc." (S5).

Note that Einstein does not use the term "mass" when he says: "[acceleration] does not in the least depend either on the material or on the physical state of the body." This leads S2 to the following conception:

"In case of gravitational field, if two bodies have the same mass, acceleration will be the same no matter what the state of the body is. [In contrast], in case of electric and magnetic fields, even if the charge is the same on two bodies, acceleration can be different." (S2). This student seems to emphasize independence of acceleration due to gravity on the state of the body (by which he perhaps means physical state and constitution) for a given mass, which is fine; but he obviously misses the key point, namely its independence on mass.

(b) and (c) The equality of inertial and gravitational masses

Three students (S1, S3, S5) had the standard conception that the equality $m_I=m_G$ was not *a priori* obvious.

"No, there is no [a priori] connection between inertial mass and gravitational mass. Their equality is not obvious. It is only by experience that we can arrive at this conclusion." (S5).

S5, while elaborating on the above point, revealed an interesting alternative conception:

"An inertial mass responds to any force. If some force acts on a body, its inertial mass (m_I) will resist it. In contrast, gravitational mass responds to gravitational force only. If a body experiences a gravitational force due to some massive object around, its gravitational mass (m_G) will resist it. But it is not at all obvious why these two things should be identical." (S5). Note the student is regarding gravitational mass as something that "resists" gravitational force (much as inertial mass resists any force). A very similar view was echoed by S4, but it led him to conclude just the opposite: the two kinds of masses must be "*a priori* equal."

"Inertial mass is a characteristic constant of a body under any form of acceleration; gravitational mass is a characteristic constant of a body under an acceleration caused by a gravitational force. So gravitational mass is a subset of inertial mass and hence they are a priori equal." (S4).

With these ideas, it is clear that none of the students really appreciated the insightful conceptual synthesis of "inertia" and "weight" that Einstein had accomplished. Three students (S1, S3, S5) rephrased the underlined quote above, but it is doubtful if they saw it as anything more than an interesting equality that Einstein made use of. The interviews reconfirmed the impressions. The students (S1, S3, S5) continued to say that $m_I=m_G$ was not *a priori* obvious while (S2, S4) said otherwise as before. S6 also claimed that it was obvious, but from earlier experience, it seemed that for S6, obvious did not mean logically self evident, but rather "obvious from experience."

2. Commentary

Most students appreciate that the universality of acceleration is characteristic of a gravitational field, in contrast to electric and magnetic fields. However, a conception of gravitational mass is shared by some of them, namely that it is the quantity that "resists" gravitational force. The standard conception of a physicist (gravitational mass is to gravitational field what charge is to electric field; both, *a priori*, unconnected to inertial mass that resists any force via the law of motion; but experimentally $m_I=m_G$) is not as easy as one might think. The alternative conception regarding gravitational mass is, in our view, a significant finding of this study.

Instruction must alert the students that the gravitational mass of a body is not to be regarded as "an inertial mass in the context of gravitational force." It is to be viewed as the gravitational analog of say electric charge. (It could well be called "gravitational charge.") The inertial mass, on the other hand, is that characteristic of the body which determines the acceleration of a body under any force, whatever its origin. The universal proportionality between the inertial mass and the gravitational mass is not *a priori* obvious. It follows from the experimentally verified fact that the acceleration of a body in a given gravitational field is independent of the material and physical state of the body.

D. Principle of Equivalence as an argument for the general principle of relativity

Einstein now introduces the Principle of Equivalence from his well known thought experiment. A man in a spacious chest in empty space, which is being accelerated "upward," will have the same experience as when he stands in a room of a house on the Earth. His legs will take up the pressure caused by acceleration and a body released from his hand will fall to the floor with the same acceleration, no matter what kind of a body it is. Since he knows the property of gravitational field (theme **C** above), he can conclude that the chest is at rest in a gravitational field acting "downward." This is the Principle of Equivalence, though Einstein never used this term in the text. The absoluteness of acceleration no longer holds and the conflict with the general principle of relativity (theme **B**) disappears. "We have thus good grounds for extending the principle of relativity to include bodies of reference which are accelerated with respect to each other, and as a result we have gained a powerful argument for the general principle of relativity." [a]

Einstein next turns the argument around to arrive at the equality $m_I=m_G$. For this he considers a body suspended in the chest "vertically" and explains the tension (T) in the rope from the point of view of both the man in the chest and an observer poised freely in space. For the former $T=m_G a$, for the latter, $T=m_I a$ where a is the acceleration of the chest for the outside observer and the gravitational field intensity for the man inside. Since both observers are equally right, the equality follows. "Guided by this example, we see that our extension of the principle of relativity implies the necessity of the law of equality of inertial and gravitational mass. Thus we have obtained a physical interpretation of the law." [b]

Einstein now revisits the example of decelerated carriage (theme **B**) which seemed to conflict with the principle of general relativity. Equipped with the new insight, he says that the carriage could be equally well regarded as at rest with respect to which there exists (during the period of application of brakes) a time-varying gravitational field in the forward direction. Einstein cautions that all gravitational fields (e.g., the Earth's field) are not of the type for which you can find another reference body with respect to which the field disappears, as in the chest example.

Students were asked to explain the underlined quotes (a) and (b) above, to explain the alternative interpretation of the

"jerk" experienced by the observer in the decelerated carriage (c), and to explain Einstein's caution above (d).

1. Students' Ideas

a. Underlined quote (a).

Only two students (S1 and S5) showed some understanding how the Principle of Equivalence helps extend the principle of relativity to accelerated frames. They made the crucial point that the absoluteness of acceleration disappears in view of this equivalence.

"Earlier we concluded that a noninertial frame has an absolute form of motion [acceleration]. But from the (chest) example it is clear that we can treat the reference body at rest and regard the forces [pseudoforces] as produced by a uniform gravitational field." (S5).

"Thus here the 'real nonuniform motion' of the carriage disappears and is accounted for by the gravity field. In this way the principle of relativity can be generalized to include accelerated frames." (S1).

The other students could not handle the question in the written responses. In the interviews, however, all of them knew that a noninertial frame is equivalent to an inertial frame with gravity and this helped Einstein generalize the principle of relativity.

b. Underlined quote (b).

Two students (S1 and S4) understood the argument leading to the equality $m_I = m_G$. S5 explained the necessity of the equality by saying that if $m_I \neq m_G$, the Principle of Equivalence would not hold. The other two students (S2 and S3) were unable to handle the question.

The interviews tested the same point by asking the students to read the article again and explain the quote (b). Everyone, including S6, continued to insist on the straightforward reasoning that since $m_I = m_G$, Principle of Equivalence is true. Despite repeatedly prodding them to reverse the reasoning, it was clear that none of them really appreciated that Einstein was turning the argument around to say that the Principle of Equivalence (and hence the general principle of relativity) *implied* the equality $m_I = m_G$. Thus the general principle of relativity interprets the equality physically, which before was a mere coincidence. Students missed Einstein's subtle reversal of the argument and, therefore, wondered at the meaning of "interpreting" the equality.

c. The decelerated carriage observer using the Principle of Equivalence.

Three students (S1, S2, S5) explained the carriage observer's view along standard lines, using the Principle of Equivalence. Two students (S3, S4) however, revealed an alternative conception:

"… he concludes that during that time, due to a varying gravitational field, all other objects (not in the carriage) such as the embankment, Earth, etc. move nonuniformly…" (emphasis ours) (S3).

"The observer can argue that there is a (nonuniform) gravitational field in the direction ahead and the jerk he experienced was due to the fact that the force of attraction [i.e., gravitational field] is different for them and the embankment." (S4).

Quite apart from the fact that the students could not explain the jerk (which arises because the upper body experiences gravitational field forward, while the same force on the feet is annulled by friction), they seemed to restrict the spatial domain of the gravitational field. According to the Principle of Equivalence, the gravitational field in the forward direction in this case is time varying but spatially uniform i.e., it is the same everywhere. This last point was missed both by S3 and S4. Interestingly, the notion of restricting the gravitational field to "inside" the reference body was also revealed by S2 in the earlier "chest" example:

"The chamber would develop gravitational field inside the chamber *only*, not outside." (emphasis ours) (S2)

This is a significant finding that agrees with the earlier finding [18] that students tend to restrict reference frames to the spatial boundaries of the reference bodies and have not internalized Einstein's notion (in the special relativity part of the text) of spatially extending the rigid body for description of phenomena with respect to a frame.

d. What misconception has Einstein warned against?

All students (except S2) understood Einstein's caution. Two of them (S1, S3) clearly revealed a feeling that a field that can be transformed away by going to another reference frame is not real.

"One may start regarding [from the chest example] gravitational fields as always apparent. That is, [one may feel] that one can always choose a reference body (with acceleration) in which no field is present. But this is not true. The Earth's field is not so." (S1)

This is quite in accordance with the standard conception that an arbitrary nonuniform gravitational field cannot be transformed away by choosing a suitable reference frame (or a coordinate system). However, Einstein did not quite like the idea of regarding the gravitational fields that can be transformed away as apparent (see [19])

2. Commentary

Students do understand that the Principle of Equivalence arises from the equality $m_I = m_G$ and some even grasp how it helped Einstein overcome the absoluteness of nonuniform motion, and thus generalize the principle of relativity. They are aware of the fact that not all gravitational fields can be made to disappear by the choice of a suitable reference body, and they regard the fields that can be transformed away as apparent. There is a tendency to restrict the domain of gravitational fields to "inside" the reference body in the usual examples of the Principle of Equivalence.

Instruction should alert the students that in the familiar example of an accelerated cabin in free space, the equivalent gravitational field exists not just "inside the cabin" but everywhere. Also, Einstein's subtle reversal of argument (using the above principle to explain the equality $m_I = m_G$) is worth highlighting here.

E. Fundamentally unsatisfactory feature of Newtonian mechanics and special relativity

Einstein points out a fundamental unsatisfactory feature of Newtonian mechanics and special relativity. They are

based on the empirical fact that with respect to only certain reference bodies (called inertial frames) in uniform translation motion with respect to each other, the law of inertia (First Law of Motion) holds well, while it does not hold for other reference bodies (called noninertial frames). Einstein asked "how does it come that certain reference bodies (or their states of motion) are given priority over other reference bodies (or their states of motion)? What is the reason for this preference?" [a] Einstein goes on to state that "the objection is of importance more especially when the state of motion of the reference body does not require any external agency for maintenance, e.g. in the case when the reference body is rotating uniformly." [b] The objection is resolved only by having a theory that holds for every body of reference, that is, by a theory consistent with the general principle of relativity. Students were asked to explain the quotes (a) and (b) above and state how the general principle of relativity resolves the problem.

1. Students' ideas

a. Reason for the preference of inertial frames (quote (a).)

All students knew that Newtonian mechanics gives preference to inertial frames and that other forces (they meant pseudoforces) arise in noninertial frames.

"The priority is due to the law of inertia." (S3). "In rotating frames, we need to put extra forces to explain the motion of a body." (S2).

But what makes some frames inertial and some noninertial?

"Classical mechanics is silent on this point. It doesn't explain why some frames are inertial and some noninertial." (S5)

This is precisely what had bothered Einstein: why some frames turn out empirically to be inertial and others noninertial. To explain the point, Einstein gives the example of two pans; one is emitting steam and the other is not. The different behavior will perplex you until you notice a "bluish something" under the pans. Analogously, the preference for inertial frames in nature is perplexing since we cannot see any *a priori* physical reason for the preference. A careful reading of the written responses suggests that students did not grasp this basic reason for Einstein's unease.

b. Why is the objection more important for a uniformly rotating reference body (quote (b))?

No student (except S1) could respond to this question adequately, and though S4 and S5 came close to handling it, they essentially repeated what was said in the text.

"Because it is clear no external force or torque acts on the uniformly rotating body. And thus it should be equivalent to other (inertial) frames. That 'something' is clearly absent here." (S1) [He is correctly referring to the same words used by Einstein in his example] "There is no other physical cause to which the invalidation of the laws of nature (in the rotating frame) can be ascribed." (S4)

"There is no external force. Without any external forces, it is not clear why things should behave differently for uniformly rotating frames." (S5)

c. How does the general principle of relativity resolve the problem of preference?

All students, of course, correctly recited that the general principle of relativity treats all frames on an equal footing and therefore resolves the problem—there are no preferred frames.

The interviews confirmed the impressions about S4 and S5 regarding question (b). S3 could not handle the question in the written responses or in the interview. S6 put it reasonably well in the interview:

"I think he is saying that for linearly accelerated frames, some force is applied and so our laws change (pseudoforces appear). But for (uniformly) rotating bodies, there is no external force, so how then are the laws different from uniformly moving frames? I think he is looking for something that can distinguish between the rotating and nonrotating frames that will tell him why it [body on a rotating frame] is feeling those pseudoforces. Newton's answer is that there is some absolute inertial frame with respect to which you can see linear acceleration or rotation and infer why the pseudo forces arise." (S6)

2. Commentary

Students clearly knew that classical mechanics privileged inertial frames over noninertial frames since in the latter the First Law was not valid and the Second Law included pseudoforces. But they only vaguely appreciated Einstein's unease in this regard. Only one student realized that for a Newtonian, there did exist an explanation for this difference, one that involved absolute space. (Inertial frames have uniform motion with respect to the absolute space; frames with nonuniform motion relative to it are noninertial). But with absolute space (ether) banished in special relativity, Einstein was left with no explanation at that stage why some frames are inertial and some noninertial.

Instruction must clearly separate the two issues: (a) in classical mechanics and special relativity, some frames (inertial frames) are singled out, and (b) there is no *a priori* physical reason in special relativity why a particular frame should be singled out i.e., turn out to be inertial. To Einstein, (a) by itself was not problematic. The real difficulty was (b). It is mysterious why of two reference bodies, neither of which needs any external agency for its motion (say two spheres rotating uniformly with respect to each other— Einstein's famous example), one should be inertial and the other noninertial. We need to emphasize that it was not the existence of difference between inertial and noninertial frames but the absence of any *a priori* physical reason for that difference, which strongly motivated Einstein to adopt the general principle of relativity.

F. Bending of light under gravity

Einstein considers two frames of reference: an inertial frame K and a frame K′ that is uniformly accelerated with respect to K. A body moving uniformly in a straight line relative to K will in general move along a curvilinear path relative to K′. Using the Principle of Equivalence, we conclude that a moving body in general follows a curvilinear path in a gravitational field. This result, however, is not new since this effect of gravity on motion is well known. "How-

ever, we obtain a new result of fundamental importance when we carry out the analogous consideration for a ray of light." Light traveling rectilinearly with respect to K will be bent in general with respect to the accelerated frame K′. If the Principle of Equivalence is regarded as valid for all phenomena, K′ may be regarded at rest with a gravitational field in a direction opposite to acceleration, and we can conclude that "in general, rays of light are propagated curvilinearly in gravitational fields." This result is a departure from Newtonian mechanics wherein gravity does not influence rectilinear propagation of light. Further it also implies that the speed of light is not a constant but depends on the position in a gravitational field. The nonconstancy of the speed of light violates one of the basic postulates of special relativity, which now becomes a limiting case of general relativity for zero gravity.

Students were asked to explain Einstein's argument above for the bending of light under gravity. The interview explored this point again, including the idea of the nonconstancy of speed of light in a gravitational field.

1. Students' ideas

a. Why is light bent under gravity?

Most students (except S5) simply repeated the argument in the text and it was not clear if they really understood it. The response from S5 was clearer:

"But when this logic [he is referring to the argument based on the Principle of Equivalence leading to curvilinear motion of material bodies under gravity] is applied to light rays, it becomes clear that it will be affected by gravity, since light rays take curvilinear paths with respect to an accelerated frame, and gravity is equivalent to acceleration." (S5).

For this theme, the interviews were more revealing. Students (except S3) tended to invoke the argument: "light has mass, so it should bend under gravity." The notion "light has mass" is acquired from their prior knowledge of special relativity ($E = mc^2$) and students evidently find it easier to use it to understand bending of light phenomenon, than to use the Principle of Equivalence for the purpose. Indeed one student (S5) went so far to say that Newtonian mechanics also predicted bending of light; he was invoking Newton's corpuscular picture of light, not realizing that this was not an accepted part of Newtonian mechanics.

Some students tended to use technical jargon (they had learnt in the course) like "null geodesics" or "space-time curvature" to explain the bending of light and had to be coaxed into the Principle of Equivalence based explanation. Also few appreciated that Einstein needed to extend the Principle of Equivalence from the restricted domain of mechanics (the so called 'weak' version of the Principle of Equivalence) to all domains of physics (the so called "strong" version of the Principle of Equivalence) to predict the bending of light.

b. Is the speed of light a constant in a gravitational field?

This question in the interviews threw up an interesting finding. Five of the six students categorically said in no uncertain terms that the speed of light does not change under gravity. S1's response was typical:

I: "What happens to the constancy of the velocity of light when it bends under gravity?"

S1: "It changes its direction."
I: "Does the speed remain the same?"
S1: "Magnitude of velocity remains the same, yes"
I: "Why do you say that?"
S1: "Because c is a fundamental constant."
I: "So the second postulate of special relativity about the constancy of speed of light remains valid here too?"
S1: "Yes, it does."

S2, S3 and S6 had similar views. S4 went further than that to state that even the velocity of light is unchanged during bending, in view of the new idea he had learnt during the course: geodesic in a curved space time.

I: "But you know light bends under gravity. How can it have a constant velocity?"

S4: "No, bending is due to the curvature of space-time. You see space-time itself is curved. So when light is moving in that curved thing, it is actually traveling straight in it."

I: "So does light travel along a straight line with constant speed in gravity?"

S4: "Yes, I mean along a geodesic."
I: "Is the speed of light also a constant?"
S4: "Yes."

The only exception was S5 who responded along standard lines (both magnitude and direction of the velocity change), but became guarded after further probing.

It is clear that the velocity of light is a major problematic point for students even after a full course on general relativity. In the route to the final formulation of general relativity, the dependence of the velocity of light on gravitational potential was a crucial early insight for Einstein, when he predicted the bending of light using the Principle of Equivalence in 1911. In the earliest scalar theory of gravity, this variable velocity of light had served the role of the gravitational potential. However, by the time he arrived at the final version in 1915, the general Gaussian coordinates had lost their metrical significance, and the question of the velocity of light became in a sense 'incommensurate' with the new theory. The modern formulation for describing light paths involves the notion of a null geodesic ($ds^2 = 0$) wherein one does not speak of velocity of light in a general coordinate system. [We can do so in a freely falling (locally inertial) frame, of course, where it continues to be c.]

Now c appears in the equations of general relativity (in nonrelativistic units) as a universal constant with the dimension of velocity and this we believe is the source of the students' response: "magnitude remains the same—direction changes during bending."

2. Commentary

For the "bending of light under gravity" phenomenon, students tend to prefer an explanation based on the notion of light carrying mass to the one based on the Principle of Equivalence. A standard course may equip them superficially with additional explanations based on the notion of light following curved paths (null geodesics) because of the space-time curvature due to gravity. However, technical exposure does not seem to guard them against the alternative conception regarding the "velocity of light in a gravitational field."

Instruction clearly needs to address the question of the velocity of light in the general theory of relativity. The question may be strictly incommensurable in this theory, yet students have not expunged their intuitive need for this concept. The "constancy of the speed of light" is such a strong "hangover" from special relativity that students tend to regard the phenomenon of gravitational bending of light as one in which the speed of light remains fixed while its direction changes due to gravity. This conception seems to get reinforced because the general relativistic equations (in nonrelativistic units) contain the universal constant c. We need to alert the students that their notion (speed of light remains fixed equal to c, but the direction of velocity changes) refers actually to the phenomenon of aberration in special relativity. For the notion of velocity of light in gravity, we may give the rough intuitive picture that the speed of light depends on the gravitational potential which varies from point to point; it is this continuous change in speed that causes the light to bend continuously (much like light bends when passing through a medium of continuously varying refractive index and hence varying speed).

G. Non-Euclidean geometry for a rotating disk observer

For further development of the theory, Einstein turned to the interpretation of space and time measurements in a general frame of reference. To illustrate the difficulties, he considers two frames: an inertial frame (K) in a gravity free domain and, for the same domain, another body of reference (K′)—a plane circular disk rotating uniformly in its plane about an axis through its center. An observer on the disk (away from the center) will experience a centrifugal force; but by the Principle of Equivalence he can regard his frame to be at rest in a gravitational field, that increases linearly with distance from the center. "The space-time distribution of this gravitational field is of a kind that would not be possible in Newton's theory of gravitation." [a]

Einstein next argues what K will infer about rods and clocks in K′, using the results of special relativity that are valid for K. For K, a clock placed at the edge of the disk is in motion and hence will show time dilation; the one at the center (at rest with respect to K) will not show this effect. Thus, as inferred by K, clocks at rest in K′ will *not* run at the same rate, the clock at the edge runs slower than the one at the center of the disk. "For this reason, it is not possible to obtain a reasonable definition of time with the aid of clocks arranged at rest with respect to the body of reference [K′]." [b]

Analogous difficulties exist for space measurements. As judged from K, a measuring rod placed tangentially along the rotating disk will contract, while a rod placed radially will not, since it is transverse to the motion. If K′ attributes unit length to all the rods in every position and orientation, he will find the circumference to diameter ratio to be greater than π, whereas for K the ratio is π. "This proves that the propositions of Euclidean geometry cannot hold exactly on a rotating disc." [c]

Students were asked to explain the quotes (a), (b) and (c) above. The interviews dwelt on (b) and (c) more probingly.

1. Students' ideas

(a) Non-Newtonian configuration of gravitational field in K′

This was not a problematic point. Most students seemed to realize that a field that was zero at the center and increased with distance did not seem possible from Newton's theory.

"Nature of this field is very different from the field arising in Newton's theory from material bodies." (S5)

S4, however, seemed to think that despite this, a gravitational field was more physical than a pseudoforce.

"This gravitational field may not be compatible with Newton's gravitational field, but it does provide physical origin of the pseudo forces." (S4)

(b) No reasonable definition of time on a rotating reference body

The argument that clocks at rest in K′ run at different rates was clear to all the students, since they were familiar with the time dilation of moving clocks in special relativity. Except for S3, they recognized its problematic implication: there was then no unique time in frame K′. Two students (S1 and S2) revealed an interesting alternative conception that clocks along a concentric circle drawn on the disk would be synchronized, but they would not be so with clocks on a different circle. Since the synchronization of clocks on a rotating disk is a technically involved matter, we do not pursue this point further here.

(c) Rotating disk and non-Euclidean geometry

In the written responses, all students (except S4) repeated Einstein's argument; the interviews helped probe the point better. Three students (S2, S3, S6) revealed the expected conception:

"… the circumference will get contracted but the radius will not, so the ratio of circumference to diameter will be less than π." (S3).

The remaining students (S1, S4, S5) argued along the standard lines that the measuring rods along the circumference would get contracted, so the ratio in the frame K′, as judged by K, would be greater than π.

S1: "The radius will remain the same, but the circumference will increase. I mean K will infer that K′ will measure it to be more since the rods used by K′ will contract along the circumference."

I: "But like the rods, the circumference of the disc should also get contracted, and so its measure should be the same."

S1: "No, circumference should be fixed for both. But rods would contract for K′. Hence the ratio (circumference/diameter) for the rotating observer, as inferred by K, would increase."

But S4 and S5 were not able to handle the question "would the circumference not contract like the rods?"

This issue about the circumference itself contracting is widely known. Einstein had tried to clarify the argument in response to queries about this point, though he does not seem to have given a detailed exposition of the rotating disk problem. One important caution is not to consider the disk being set to rotation from rest but to see it as already rotating (see [20]). The standard argument seems to be as follows: Imagine a circular ring at rest in K that just surrounds the rotating disk. The circumference to diameter ratio of the ring is π for

K. By symmetry, K′ should find the ring to be a circle and conclude that the circumference of the disk equals that of the ring, since they both overlap. So the circumference of the disk remains the same. But the measuring rods in K′ along the circumference contract, so K concludes that K′ would need a greater number of rods to span the circumference of the disk than he (K) would to span the ring; hence the result.

The argument assumes that K′ takes the length of a standard infinitesimal rigid rod to be the same everywhere. None of the students showed explicit awareness of this basic assumption.

2. Commentary

Students are able to see that the gravitational field corresponding to a rotating disk is not like an ordinary Newtonian field. They appreciate that it is not possible to define a unique time in the rotating frame as a whole, since clocks at rest in the frame run at different rates. Uncritical use of prior knowledge of special relativity, however, leads them to conclude that synchronization of clocks is possible along concentric circles, but not across these circles. The rotating disk geometry presents the expected difficulty, namely, that students do not see why the circumference of the disk should not contract like the measuring rods used to measure its length.

Instruction should alert the students that the clocks on the rim of a rotating disk cannot be synchronized, though they go at the same dilated rates for the outside inertial observer. This is because two inertial observers instantaneously at rest with respect to two different points on the rim have relative motion; so according to special relativity, the mutual synchronization of their clocks is not possible. We also need to take them away from the notion that an observer can arrive at the laws of physics only in his frame and that he can say nothing about the laws in another frame. Students' difficulties in this theme partly arise from the fact that it is the outside inertial observer who infers about the problems (lack of synchronization of clocks, non-Euclidean character, etc.) of the rotating disk observer.

H. Simple illustration of a non-Euclidean continuum

Einstein next elucidates the notion of a non-Euclidean continuum by a simple example. Consider the surface of a plane marble table. Since we can go from one point to another of the surface continuously without "jumps," the surface is a continuum. Take little rods of equal length (i.e., those which can be laid on each other exactly end to end) and begin by making a square on the surface with four of these rods. Go on adding squares one on another until the whole slab is covered. Einstein emphasizes that it is not logically obvious that the construction would succeed. "If at any moment three squares meet at a corner, then two sides of the fourth square are already laid, and, as a consequence, the arrangement of the remaining two sides of the square is already determined. But I am no longer able to adjust the quadrilateral so that its diagonals may be equal." [a] The fact that these turn out to be equal is a property of the slab and the rods: For this situation (Case I), Euclidean geometry is

valid and we can set up the usual Cartesian coordinate system to specify any point on the slab.

Next (Case II), suppose we heat the table differentially, say more at the center than at the periphery, so the temperature is nonuniform over the slab. Suppose the rods expand, the increase in length being proportional to the increase in temperature. The "square mesh construction" of Case I will now fail. However, we could still continue to regard the slab as a Euclidean continuum. This is because there might exist rods of a special material which are not influenced by temperature, using which the "square mesh construction" would still succeed. Thus in this case the failure of the "square mesh construction" with the expandable little rods is attributed to the varying lengths of those rods at different positions, and not to any basically different property of the continuum itself.

But now imagine the situation (Case III) where the rods of every material expand identically with temperature and there is no other way of detecting the effect of temperature. In this case, the "square mesh construction" would fail for every kind of rod. We must naturally assign the same length (say unit length) to every little rod no matter where it is placed on the differentially heated marble table surface. With this assignment, Euclidean geometry would be violated on the slab. In this case then, the slab is a non-Euclidean continuum.

Students were asked what Einstein was trying to say in quote (a); why in Case II we can continue to regard the surface as a Euclidean continuum and why in Case III we should regard it as a non-Euclidean continuum.

1. Students' Ideas

The written responses show that students missed the basic point of the example—namely that it is the measurements that determine whether the continuum is Euclidean or otherwise. Four students understood the simple property of the "square-mesh" construction and repeated the assertion in the text that its success shows the Euclidean nature of the slab. S1 and S5 put it better than the other two (S2 and S3):

"If we find that the whole slab has been covered with perfectly symmetric squares with equal diagonals and the rules of Euclidean geometry are completely followed by the square mesh just laid, we can say that the marble slab constitutes a Euclidean continuum." (S5)

But none of them explicitly stated that this success is not *a priori* obvious but has to be established by measurements (in this case by little rods of unit length). A possible inference is that for most students, Euclidean geometry is "obvious." This inference is likely to apply to S4 also, who did not even consider Case I for comments.

Case II revealed a fundamental alternative conception. Three students (S2, S3, S4) surmise that the slab continues to be a Euclidean continuum despite the differential heating, since though the rods expand, the table does not. S4 put it more explicitly saying that this would cause a bulge at the center.

"We can maintain this point of view [Euclidean nature of the continuum for Case II] by keeping in mind that the coefficient of expansion of marble is nearly zero. So whatever may happen to the rod mesh—they may bulge in or out, the

marble slab remains a plane Euclidean continuum" [S4]

S1 and S5 repeated the text assertion that for Case II, the continuum is Euclidean because we can have other measuring tools that are unaffected by temperature.

"We can choose rods of some other materials which are not affected by temperature." (S5)

S1 even suggested a different measurement tool unaffected by temperature:

"Make measurements using light, photo detectors and good clocks… the marble space will continue to be a Euclidean continuum." (S1)

The alternative conception revealed by S4 above regards the continuum itself as a physical entity which expands or does not expand with temperature and whose properties can be ascertained per se. This is to be contrasted with the standard conception: the continuum does not exist independent of measurements; its geometrical properties can be ascertained only through measurements by means of some agreed standard tools (e.g., rods of negligible coefficient of expansion in Case II). The importance of this point for conceptual understanding of the curved space-time continuum of general relativity can hardly be overemphasized (see [21]).

None of the students showed a clear understanding of Case III. The responses of S2 were rather muddled and not amenable to a clear analysis. S4 continued to think in terms of expansion of *both* the marble table and the rods and arrived at the conclusion that contradicted the text:

"If on heating, all the materials including the rod mesh and the marble table bulge to the same degree, then even on top of the bulge of the marble slab, the rod mesh has the shape of squares." (S4)

S3 was not able to explain the assertion in the text but instead came to a different conclusion.

"Indirectly, it says that it is not possible to measure distances in a non-Euclidean continuum." (S3)

This feeling (see also S5's response below) is related to students' unease regarding distance measurements on a curved surface; the straight rigid rod would not do, they all need to be bent; then how do you measure distance?

S1 and S5 repeated Einstein's argument regarding Case III.

"However, if every rod behaves in the same manner, we are forced to accept that the surface is non-Euclidean." (S5)

S5, however, could not get rid of the embedding picture—viewing a continuum from the point of view of a higher dimensional space in which it is embedded. When asked why we cannot use Cartesian coordinates on a sphere, he said:

"On the surface of a sphere, we cannot use the same rods which are used to describe the surface of a marble slab. We are forced to use the rods which can be laid out on the surface of the sphere and these rods will be naturally of a very different nature [he means that they will be curved] than those used for a surface of the Euclidean type." (S5)

The embedding tendency came out even more clearly in the interviews (see below).

An important point in this example related to what length to assign to the little measuring rods at different positions and orientations of the differentially heated marble table. Clearly, for Case II, we must assign different lengths at different locations of the rods that expand by amounts that vary with temperature. We can do so since we can compare these lengths with those of the standard little rods that do not expand with temperature. The marble slab can still be regarded as a Euclidean continuum, since though the square-mesh construction will fail for the rods that expand, it will succeed with those that do not expand. But what about Case III, wherein all rods of whatever material expand identically? What lengths should we now assign to the rods at different positions or orientations? Einstein argues that we must assign the same length (say unit length) to each rod, whatever its location, for we can do nothing else: any other assignment will be completely arbitrary. With unit length to every rod and the square mesh construction failing, we must conclude that the continuum is non-Euclidean. Students did not dwell on this issue in their written responses, so in the interviews we probed it in some detail. The interviews were indeed revealing, regarding the question of what length to assign to the rods at different locations in Case III.

S2's responses were muddled and led to nothing significant. S1 thought that there was nothing like unity [unit length] defined in a non-Euclidean continuum.

I: "For Case III, what length do we assign to each rod at different places?"

S1: "The lengths won't be the same."

I: "But Einstein says that we should assign unit length to each rod everywhere."

S1: "That we can do."

I: "Does it make sense? Will it not mean giving different distances the same unit value?"

S1: "There is nothing like unit length defined since the space is non-Euclidean."

S3 and S6 said more explicitly that we must assign different lengths at different locations even in Case III. S4, though otherwise unclear, came close to grasping the essence.

I: "So, should we assign unit lengths to all rods in Case III?"

S4: "Yes, because he is saying that we have nowhere else to go and check."

S5 continued to use the embedding picture that he had revealed in the written responses.

I: "So, are you going to assign unit lengths to the rods everywhere? Or will you assign different lengths to rods at different places in Case III?"

S5: "No, I think we will assign 'one' to the rod length everywhere. But 'one' at different places will be different." [S1's conception]

I: "What does that mean: 'one' is different at different places? If I know it is different, I will assign a different number."

S5: "If we see the table from outside, then for us it will be different. But for an observer on that manifold, it will be the same everywhere."

I: "Why?"

S5: "Well, he cannot view it from outside and know that it is of a different length, so he must conclude that it is of unit length everywhere."

Note the unease of S5; he correctly knows the two-dimensional surface view but believes the "truer" picture is that of the observer outside (in the three-dimensional space).

2. *Commentary*

The example of the differentially heated marble table uncovers some of the most basic alternative conceptions of students regarding Euclidean and non-Euclidean geometry. First, they do not regard Euclidean character an empirical property of a continuum. Second, their visualization that "straight" rods cannot be used to measure distances on a curved surface (e.g., a sphere) leads them to doubt the ability to measure distances on a non-Euclidean continuum. Third, they find it hard to absorb the view of the observer "in" the surface, but instead go more readily for the view of the "outside" observer in the higher dimensional space in which the surface is embedded. This is what makes them uncomfortable with assigning equal lengths (unit lengths) to all rods on the differentially heated surface (in Case III above), since it seems to conflict with the "outside neutral" view. And even when they agree to assign unit lengths to all the rods everywhere, they do not regard the assignment as "true." The embedding picture persists in that the "unity" in "unit length" is regarded as not being well defined or fixed. Fourth, the continuum itself, besides the measuring rods, is regarded as a physical entity which "expands" or "bulges."

Instruction needs to emphasize that Euclidean geometry, however "obvious" it may seem, needs to be checked for a given manifold empirically by measuring distances using some standard device (e.g., rigid rods). Also on a curved manifold (e.g., a sphere) the distance between any two points *can* be measured using some agreed measurement device—the fact that "straight" rigid rods are no longer available does not preclude this. Also by measurement of such distances, the non-Euclidean character of the manifold can be established, without the need to view the manifold as embedded in a Euclidean manifold of some higher dimension. The embedding picture should not be allowed to come in the way of a deeper appreciation that the manifold may not exist as a separate entity apart from the measurements.

I. Non-Euclidean space-time continuum of general relativity

The example of rotating disk suggests that for general reference bodies in arbitrary motion, it is not possible to build a reference system of rigid rods and clocks in the manner familiar in Galilean and special relativity. To grapple with this loss of metrical meaning of coordinates, Einstein argued that every physical phenomenon is in the final analysis reducible to a series of encounters, which can be labeled by arbitrary numbers, provided they uniquely identify the encounter and satisfy the requirement of continuity. Thus any event may be assigned four arbitrary (Gaussian) coordinates (x_1, x_2, x_3, x_4); together the events form a four-dimensional continuum. The notion of "distance" between two neighboring events is no longer of the Euclidean type—the space-time is non-Euclidean. However in a sufficiently small region, the continuum may still be regarded as Euclidean. This is the geometrical analog of the physical fact that a freely falling frame in an arbitrary gravitational field is locally inertial.

With the frames of reference thus replaced by Gaussian coordinates, Einstein extends the scope of his original principle of general relativity (all frames of reference are on the same footing) to the final form—all Gaussian coordinate systems are equivalent. With the Principle of Equivalence as a heuristic tool, Einstein finally solves the twofold problem of gravitation—obtaining (a) the law of motion of a material point under a gravitational field and (b) the general law of the gravitational field. For (b), he additionally uses the requirement that the field and matter together satisfy the laws of conservation of energy and momentum.

This last theme is rather technical and the details of our study on this theme will be omitted. Here we make only two brief observations:

(1) We have to alert students against the natural feeling (seen in the study) that an event must happen at some place and time in some coordinate system and that the loss of metrical meaning of the coordinates in general relativity arises because of the arbitrary transformations of the initially space-time meaningful coordinates. This view is correct only in special situations, not in general. The special situations are important parts of a standard teaching course, which is perhaps the reason students continue to hold the view even after a course.

(2) Students are well aware of the principle of general covariance which puts all Gaussian coordinate systems on an equal footing (they even know that the language of tensors is best suited to capture this equivalence), but do not appreciate that this goes beyond treating all reference bodies on the same footing. Einstein's vital step of extending the scope of his principle (from all bodies of reference to all coordinate systems) is missed out mainly because of a lack of historical perspective in a standard course.

IV. DISCUSSION

We now discuss briefly the possible underlying sources of the various alternative conceptions of students described above:

(1) *Alternative conceptions in antecedent domains* which survive instruction reappear and affect students' learning of new domains based on them. We see notable instances of this here: regarding relativity principle as self-evident (theme **A**); confinement of gravitational fields (in the context of the Principle of Equivalence) to the spatial boundaries of reference frames (theme **D**), etc. The tendency to understand the abstract in concrete terms underlies several of the alternative conceptions in Galilean relativity found in our earlier work [22]. This reappears in the present work in students' anthropomorphic notions regarding the "observer," "difficulty in discovering laws in noninertial frames" (theme **B**) and in their difficulty to understand that an observer could infer laws in another frame (theme **G**). The most striking example of this tendency in a novel situation was seen in theme **H**, wherein students view the continuum as a concrete entity and arrive at its geometry through its physical behavior (thermal expansion) instead of looking at measurements of distances by standard tools.

(2) *Prior knowledge* influences how students try to come to terms with the novel ideas of general relativity. The notable examples here are: assimilation of the notion of gravitational mass in the already learnt notion of inertial mass (familiar to them from school physics) (theme **C**), their holding on to the idea of the constancy of speed of light even while bending under gravity (theme **F**), their preference for the "light has mass, so it bends" argument to the Principle of Equivalence based argument (theme **F**), an obvious influence of their prior learning of special relativity, which also leads them to the notion that clocks on concentric circles of a rotating disk are synchronized (theme **G**). For most of these situations, their prior knowledge is along standard lines, but is being used uncritically in a new domain.

(3) The *finer conceptual discriminations* necessary in a domain are often ignored or subsumed under the broader conceptions of students (see [23]). This may arise not merely because of inattention to the nuances and detail but may also be rooted in students' epistemologies, an aspect not probed in this paper. Examples include their not distinguishing the logically self-evident from the empirically obvious (theme **A**), their not appreciating Einstein's reversal of argument to "explain" the equality of inertial and gravitational masses, using the Principle of Equivalence (theme **D**), their lack of awareness of Einstein's basic assumption in the rotating disk example [taking the length of a standard rod to have the same length everywhere (theme **G**)], etc. This also shows up in their generally imprecise use of scientific language (e.g., not distinguishing "laws of nature" from "description of phenomena," ambiguous use of the word "obvious," etc.), a problem perhaps compounded by the fact that for the students under the present study, the medium of instruction (English) is not their first language.

We conclude by clarifying the basic motivation of this study. We are not dwelling here on the universality or frequency of the various conceptions that students show up on the basic (nontechnical) themes of general relativity. The sample is obviously too small to make any statement of that kind. Nor is there a claim of discovering any consistent alternative schemata of the few cases (students) we have studied in detail. Indeed, in both the written responses and the interviews, students' cognitive behavior was rather inconsistent, as can be expected in a domain such as general relativity. Our aim was simply to discover the alternative conceptions that arise as students attempt to learn this subject at the introductory level. Needless to say, the interpretation of students' ideas and their contrast with the standard concepts is circumscribed by our own limited understanding of this highly subtle discipline. Still, the findings of this study, we believe, should be of value to physics education researchers as well as to the teachers of the subject.

[1] S. Panse, J. Ramadas, and A. Kumar, Alternative Conceptions in Galilean Relativity-Frames of Reference, Int. J. Sci. Educ. **16**, 63 (1994).

[2] J. Ramadas, S. Barve, and A. Kumar, Alternative conceptions in Galilean relativity: Distance, time, energy and laws, Int. J. Sci. Educ. **18**, 463 (1996).

[3] J. Ramadas, S. Barve, and A. Kumar, Alternative Conceptions in Galilean Relativity: Inertial and Non-Inertial Observers, Int. J. Sci. Educ. **18**, 615 (1996).

[4] A. Bandyopadhyay, Students' ideas of the meaning of the relativity principle, Eur. J. Phys. **30**, 1239 (2009).

[5] L. McDermott and E. Redish, Resource Letter: PER-1: Physics Education Research, Am. J. Phys. **67**, 755 (1999).

[6] J. Hartle, General relativity in the undergraduate physics curriculum, Am. J. Phys. **74**, 14 (2006).

[7] R. Wald, Resource Letter TMGR-1: Teaching the mathematics of general relativity, Am. J. Phys. **74**, 471 (2006).

[8] S. Chandrasekhar, On the "Derivation" of Einstein's Field Equations, Am. J. Phys. **40**, 224 (1972).

[9] W. Rindler, General relativity before special relativity: An unconventional overview of relativity theory, Am. J. Phys. **62**, 887 (1994).

[10] A. diSessa, Toward an epistemology of physics, Cogn. Instruct. **10**, 105 (1993).

[11] S. Vosniadou, On the Nature of Naïve Physics, in *Reconsidering the Processes of Conceptual Change*, edited by M. Limon and L. Mason (Kluwer Academic Publishers, 2002), pp. 61–76.

[12] A. Einstein, *Relativity: The Special and the General Theory* (Edition, Bonanza Books, New York, 1961).

[13] R. Resnick, *Introduction to Special Relativity* (Wiley, New York, 1968).

[14] D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics* (Wiley, New York, 1996).

[15] R. D. Inverno, *Introducing Einstein's Relativity* (Oxford University Press, New York, 1992).

[16] Students' responses (italicized throughout) are not verbatim quotations. They have been edited and also rephrased where necessary for clarity, but we have taken care to see that the intended explanation or idea of the student is faithfully conveyed.

[17] See Ref. [4].

[18] See Ref. [1].

[19] J. D. Norton, What was Einstein's Principle of Equivalence?

Stud. Hist. Philos. Sci **16**, 203 (1985).

[20] J. Stachel, The Rigidly Rotating Disc as the "Missing Link" in the History of General Relativity, in *Einstein from B to Z*, (Birkhäuser, Boston, 2002).

[21] See Appendix V of Ref. [12].

[22] See Ref. [1].

[23] F. Reif, *Applying Cognitive Science to Education: Thinking and Learning in Scientific and Other Complex Domains* (MIT Press, Cambridge, MA, 2008).