Understanding curved spacetime

The role of the rubber sheet analogy in learning general relativity

Abstract

According to general relativity (GR), we live in a four-dimensional curved universe. Since the human mind cannot visualize those four dimensions, a popular analogy compares the universe to a two-dimensional rubber sheet distorted by massive objects. This analogy is often used when teaching GR to upper secondary and undergraduate physics students. However, physicists and physics educators criticize the analogy for being inaccurate and for introducing conceptual conflicts. Addressing these criticisms, we analyze the rubber sheet analogy through systematic metaphor analysis of textbooks and research literature, and present an empirical analysis of upper secondary school students' use and understanding of the analogy. Taking a theoretical perspective of embodied cognition allows us to account for the relationship between the experiential and sensory aspects of the metaphor in relation to the abstract nature of spacetime. We employ methods of metaphor and thematic analysis to study written accounts of small groups of 97 students (18-19 years old) who worked with a collaborative online learning environment as part of their regular physics lessons in five classes in Norway. Students generated conceptual metaphors found in the literature as well as novel ones that led to different conceptions of gravity than those held by experts in the field. Even though most students showed awareness of some limitations of the analogy, we observed a conflict between students’ embodied understanding of gravity and the abstract description of GR. This conflict might add to the common perception of GR being counter-intuitive. In making explicit strengths and weaknesses of the rubber sheet analogy and learners’ conceptual difficulties, our results offer guidance for teaching GR. More generally, these findings contribute to the epistemological implications of employing specific scientific metaphors in classrooms.

1. Introduction

The Earth circles around the Sun and we stay grounded on Earth because of gravity. Yet, the nature of gravity eluded human understanding for centuries. It was only with Albert Einstein’s theory of general relativity (1915) that physicists found a fundamental description for gravity which set the stage for the development of modern physics in the 20th century. General relativity (hereafter GR) is a modern theory of gravitation that extends classical mechanics to cosmic scales. Motion at the speed of light and physics close to extremely massive objects such as black holes require a more powerful framework than Newton’s classical mechanics can offer. By describing gravity as geometry, GR offers such a framework with greater explanatory power than classical mechanics: the fabric of our universe can be modeled by four-dimensional spacetime and it is the curvature of spacetime that manifests itself in form of gravity.

Gravity as a manifestation of curved spacetime is an abstract concept that students, not being able to rely on the advanced mathematical formalism, must grasp in terms of other areas of experience. Thus, the description of this concept requires metaphoric language. In this study, conducted within the Norwegian design-based research project ReleQuant, we aimed to understand how upper secondary students reason with analogies and metaphors to conceptualize gravity and curved spacetime.

The formulation of GR not only provided impetus to the further development of physics, but it also inspired the emergence of new fields such as cosmology. In fact, the significance of the theory extended beyond the mere contents of scientific laws and theories. Indeed, adopting a relativistic perspective entailed a change in the worldview of many scientists at the time of Einstein (Chandler, 1994). With its apparent metaphysical implications, GR brought about a new heyday of the philosophy of space and time as well (Reichenbach, 1928).

With Einstein's insight into the deep connection between gravity, space, and time, a century-long dispute on the nature of space and time found its culmination. The dispute reaches back in the beginning of the 18th century to Newton and Leibniz, who held opposing views on this topic (Vailati, 1997). Whereas Leibniz argued that space and time are relational and can only be defined through orderings between objects, Newton described space and time as absolute entities that are as real as any object in the world. The Newtonian view of absolute space and time dominated academic discourse for almost 200 years. It was only at the end of the 19th century that philosophers and scientists started to question notions of absolute space and time (Mach, 1893; Poincaré, 1898). These considerations were predecessors to the revolutionary ideas of Einstein who eventually replaced absolute space and time with the notion of dynamical spacetime - a replacement whose philosophical impact still can be felt today (Chandler, 1994).

Surprisingly, the great importance of GR in physics and philosophy has not corresponded to equivalent attention in education on how students understand such concepts. Even though current fields of physics research such as gravitational wave astronomy (Abbott et al., 2016) as well as the working of modern communication technologies rest greatly on our relativistic understanding of gravity, physics in high schools remains mostly dominated by classical theories of gravity (Henriksen et al., 2014; Velentzas & Halkia, 2013). However, students are confronted with a growing number of representations in the media and popular culture, such as in recent discoveries about gravitational waves, which present gravity as a relativistic phenomenon. While other domains of modern physics such as quantum physics and special relativity have already entered high school and undergraduate education in many countries (Henriksen et al., 2014; Krijtenburg-Lewerissa, Pol, Brinkman, & Van Joolingen, 2017; Levrini, 2014; Stadermann & Goedhart, 2017), it was only very recently that physics educators made first attempts to introduce GR to school curricula and to investigate students' understandings of it (Kaur, Blair, Moschilla, Stannard, & Zadnik, 2017a; Kersting, Henriksen, Bøe, & Angell, 2018). In a society that is pushing knowledge and technological advancement ever further, it is important to teach students our best understanding of the universe, and this can only be done if we know how to communicate relativistic concepts effectively.

Studies on secondary school students' conceptual development of key concepts in GR are scarce (Kersting et al., 2018). Existing research either looks at special relativity instead of general relativity (Dimitriadi & Halkia, 2012; Levrini, 2014; Levrini & DiSessa, 2008) or studies undergraduate physics learning (Bandyopadhyay & Kumar, 2010a, 2010b; Hartle, 2005). Based mostly on case studies and interviews, the findings in these studies suggest that students often struggle with the interpretation of relativistic concepts and phenomena.

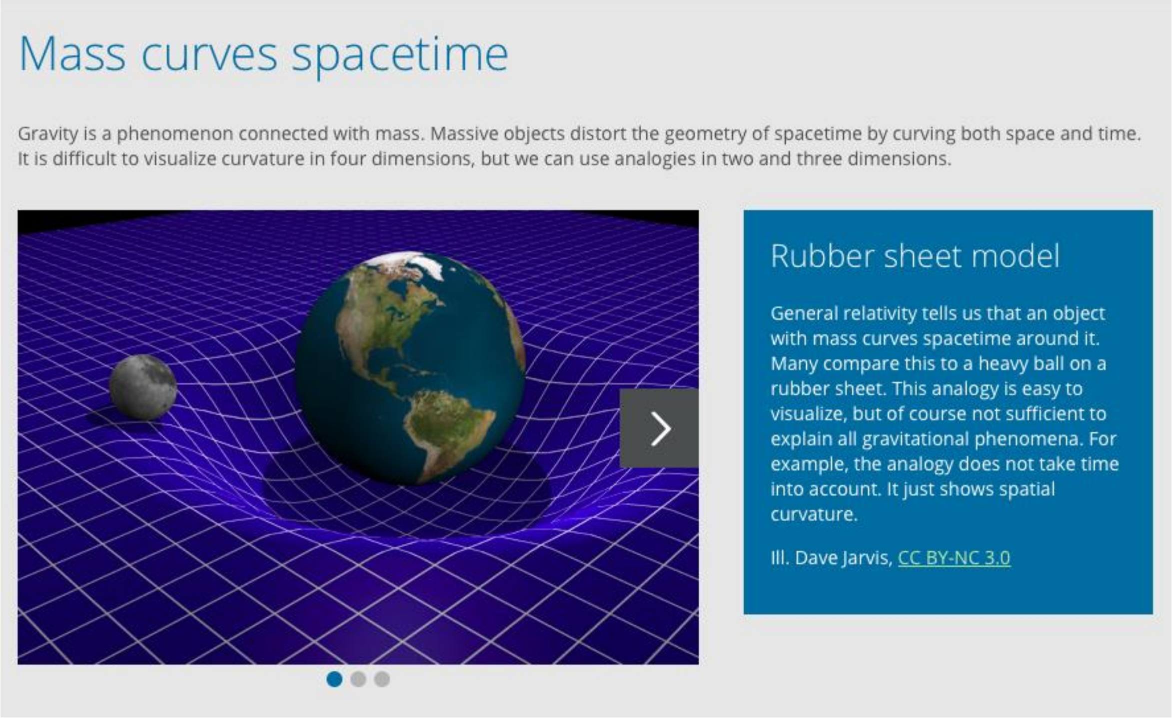
Recently, educational projects in Australia and Norway (Kaur et al., 2017a; Kaur, Blair, Moschilla, Stannard, & Zadnik, 2017b; Kersting et al., 2018) have started to investigate the learning of GR at the high school level in response to increased emphasis in national curricula. Efforts in Australia rely on so-called enrichment programs that introduce modern concepts of space and time to 10-16 year-old students. Work in Norway relies on digital learning resources that were trialled with 18-19 year-old students. In an attempt to achieve an educational reconstruction of GR, we reviewed the literature to identify the main challenges of teaching and learning relativity (Kersting et al., 2018). General challenges include the advanced level of mathematics, the lacking experience with relativistic phenomena, and the counterintuitive nature of these phenomena in light of classical physics. More specific challenges concern the role of observers in different reference frames and the Euclidean nature of our universe that students take for granted. Despite those challenges, the results from Australia and Norway are encouraging. Findings suggest that younger students are motivated by topics of Einsteinian Physics (Kaur et al., 2017b) and that students can gain a qualitative understanding of GR when provided with appropriately designed learning resources and support from peers (Kersting et al., 2018).

Moreover, the latter study is among the first to present empirical results on upper secondary students' understanding of curved spacetime. Focus group interviews revealed that spacetime is an engaging, yet challenging concept that students felt very uncertain about. The only other study that we are aware of to report on students' conceptual understanding of spacetime looked at senior undergraduate students taking a course on

GR (Bandyopadhyay & Kumar, 2010b). However, these researchers only touched upon non-Euclidean geometry and did not investigate students' understanding of the geometry of spacetime in detail. Therefore, the conceptual understanding of curved spacetime still seems to be a mostly unexplored topic in science education research.

Teaching GR on undergraduate and upper secondary school level requires teaching approaches that rely on qualitative explanations and elementary mathematics (Kersting et al., 2018). Such approaches entail the use of thought experiments (Velentzas & Halkia, 2013), geometric models (diSessa, 1981; Zahn & Kraus, 2014), hands-on experiments (Pitts, Venville, Blair, & Zadnik, 2014), and simple mathematical approximations (Stannard, Blair, Zadnik, & Kaur, 2017). Common to these teaching strategies is the shared understanding that the mathematical foundation of GR is very abstract and that many of its consequences are counterintuitive (Bandyopadhyay & Kumar, 2010b; Kersting et al., 2018). These challenges affect high school and undergraduate students alike, because GR contradicts what most students have learned in previous physics classes, namely that gravity is a force.

While there seems to be consensus about the educational challenges of GR, the most prevailing popular representation of the theory gives rise to a debate among physicists and physics educators. Both in teaching resources and in popular science culture, the so-called rubber sheet analogy (hereafter RSA) is a widely used tool to make sense of four-dimensional curved spacetime (Greene, 2010). The analogy compares the fabric of the universe to a stretched rubber sheet (Figure 1). Gravitation and the dynamic interplay between the movement of massive objects and the curvature of spacetime are illustrated by placing a bowling ball and marbles on the rubber sheet. The bowling ball produces a warp of the rubber, which results in an inward tug that will influence the movement of the marbles. The bowling ball represents for example the Earth and the marble is like the Moon circling around the massive ball. It is the warp of the rubber sheet that creates the gravitational tug. There is no need to introduce a force that, mysteriously, acts at a distance.



Mass curves spacetime

Gravity is a phenomenon connected with mass. Massive objects distort the geometry of spacetime by curving both space and time. It is difficult to visualize curvature in four dimensions, but we can use analogies in two and three dimensions.

Rubber sheet model

General relativity tells us that an object with mass curves spacetime around it. Many compare this to a heavy ball on a rubber sheet. This analogy is easy to visualize, but of course not sufficient to explain all gravitational phenomena. For example, the analogy does not take time into account. It just shows spatial curvature.

Ill. Dave Jarvis, CC BY-NC 3.0

Fig. 1 **A screenshot of the Norwegian learning environment that introduces the rubber sheet analogy.**

The ubiquity of the RSA in teaching resources and popular science literature nowadays stems from the challenge to visualize a theory whose geometry continues to confound. Einstein had admitted that our imaginative faculty cannot conceive of four dimensions:

> No man can visualize four dimensions, except mathematically. We cannot even visualize three dimensions. I think in four dimensions, but only abstractly. The human mind can picture these dimensions no more than it can envisage electricity. Nevertheless, they are no less real than electro-magnetism, the force which controls our universe, within, and by which we have our being.
> (Einstein in Viereck, 1929)

In response to this challenge, Einstein was presumably the first to employ the analogy that compares spacetime to a cloth. In a correspondence with his colleague Willem de Sitter, who would later publish joint work with Einstein on the curvature of the universe, Einstein explained: "Our problem can be illustrated with a nice analogy. I compare the space to a cloth floating (at rest) in the air, a certain part of which we can observe. This part is slightly curved similarly to a small section of a sphere's surface." (Hentschel, 1998, p. 301)

Only shortly after the publication of GR, Einstein attempted to present the theory of relativity to a more general audience (Einstein, 1917). Similar expositions by others followed shortly after that, and already in 1925 the eminent mathematician and philosopher Bertrand Russell used a “soft india-rubber” to illustrate the idea of curved spacetime (Russell, 1925). Interestingly, at the same time Russell cautioned of the risks of simplifying scientific ideas too much: “Einstein revolutionized our conception of the physical world, but the innumerable popular accounts of his theory generally cease to be intelligible at the point where they begin to say something important” - a wise remark that foreshadowed the debate around the RSA that scientists and educators still lead today.

On the one hand, advocates of the RSA have praised it as an “excellent analogy” (Thorne, 2009, p. 77) because of its visual power and its intuitive appeal both to students (Farr, Schelbert, & Trouille, 2012) and physicists in the field:

> The rubber membrane-bowling ball analogy is valuable because it gives us a visual image with which we can grasp tangibly what we mean by a warp in the spatial fabric of the universe. Physicists often use this and similar analogies to guide their own intuition regarding gravitation and curvature. (Greene, 2010, p. 71)

On the other hand, critics consider the RSA to be “misleading” (Price, 2016, p. 588) and to pose a “considerable risk to the formation of misconceptions” among students (Zahn & Kraus, 2014), because of oversimplification and incorrect presentation of the physics:

> Unfortunately, the illustration makes no sense. Students observe that space is not a rubber sheet, does not curve into an unseen dimension, and does not push objects into circular orbits. The rubber sheet does not even reflect the symmetry of the central mass—if you turn the illustration upside down the explanation fails. (Gould, 2016, p. 396)

Seeing that experts hold divided opinions on the educational value of the RSA when teaching GR, it is surprising that the ongoing debate is mostly based on opinions and claims without a proper evidential base. Gould, for example, claimed that “ (…) students are often confused by literal illustrations of the concept [of curved spacetime]” (2016, p. 396), but he presented no empirical evidence to support this claim. Looking into the literature, the works that address the RSA explicitly can be grouped into two camps. On one side, physicists focus on the mathematics of the RSA to show why the analogy can be an instructive teaching tool (Middleton & Weller, 2016), or to replace it with more appropriate mathematical models (Gould, 2016; Price, 2016). On the other side, science educators investigate how students understand the RSA (Baldy, 2007; Steier & Kersting, n.d.; Watkins, 2014). However, these very few investigations have, so far, addressed the RSA rather as a way to explain gravitational phenomena in the framework of Newtonian physics rather than to shed light onto how the RSA might facilitate students’ understanding of curved spacetime in the context of GR.

Addressing the problem that secondary students display with a force of gravity that acts magically at distance, Baldy (2007) introduced the “pillow-model” to study French ninth-graders’ (15 years old) ideas of attraction between objects. The pillow-model replaces the rubber sheet by a soft pillow, but serves conceptually the same purpose as the RSA. Baldy compared two teaching methods, one based on Newtonian physics and one based on the pillow-model, and studied student’s conceptions of falling bodies. She found that the Newtonian approach is less effective, even though she admitted that her results "(…) are not intended to mean that the students built a representation of the universe that conformed to Einstein's theory on all points, nor that they understood the theory." (Baldy, 2007, p. 1784)

In an exploratory study on the conceptual understanding of curved spacetime that was conducted within the same project as the present work, we analyzed a discussion between two Norwegian upper secondary physics students who showed deep engagement with gravity and spacetime, but struggled to accept certain aspects of the new concepts. The results suggested that the RSA might be problematic for learners, because it makes use of two different concepts of gravity and relies on classical gravity to make the analogy of "Einsteinian" gravity work. (The fact that the analogy draws on classical gravity, like the force that creates a well in the rubber sheet, to explain a new interpretation of gravity lets the pair of students struggle conceptually.) However, it is not clear whether these results can be readily generalized to a broader sample of students (Steier & Kersting, n.d.) .

Addressing the controversy around the use of the RSA in the domain of GR, we want to bring the debate forward by offering actual empirical results on upper secondary school students' ideas about curved spacetime in relation to the RSA. Insights into students' understanding and their use of the most common representation of curved spacetime are critical in order to investigate learning processes and conceptualization of spacetime in GR and to develop efficient teaching approaches.

We aim to understand how upper secondary students reason with the RSA to conceptualize gravity and curved spacetime. To guide our examination, we ask the following research questions:

1. What features of gravity as they were explained by Einstein does the rubber sheet analogy hide and highlight?
2. What characterizes students' understanding of the rubber sheet analogy?
3. In what ways do students show awareness of the analogical nature of the rubber sheet analogy when conceptualizing gravity and curved spacetime?

We hope that addressing these questions will serve as an impetus for the ongoing educational debate around the RSA and that it will add to the emerging body of knowledge concerning the teaching GR. More generally, we hope that our findings will contribute to the epistemological implications of employing specific scientific metaphors and analogies in science classrooms.

2. Theoretical Background

In the following sections, we frame the challenge of analyzing the RSA and students' ideas of curved spacetime in relation to research about the use of analogies and metaphors in science education.

2.1. Analogies, Metaphors, and Embodied Cognition in Science Education

Both the wish to approach GR from a qualitative perspective and our inability to visualize four dimensions make the RSA an appealing tool to communicate aspects of curved spacetime. Indeed, instructional analogies and metaphors have become a popular tool in science education, because they can help to communicate abstract scientific concepts (Aubusson, Harrison, & Ritchie, 2006). However, science educators have also recognized limitations to this approach due to the often-unpredictable ways that students interpret analogies and metaphors (Harrison & Treagust, 2006).

Before we unpack further aspects of this criticism in relation to the RSA, let us define what we mean by an analogy or a metaphor in our context. Niebert et al. (2012) reviewed the use of both terms in the science education literature and came to the conclusion that most science educators treat analogies and metaphors synonymously as statements that characterize one thing in terms of another. This characterization goes back to Lakoff and Johnson whose broad definition of metaphors encompasses analogies as well (Lakoff & Johnson, 2003). Genter et al. observed that the processes of understanding metaphors and analogies are the same (Gentner, Bowdle, Wolff, & Boronat, 2001). On the basis of this observation, Niebert et al. concluded that the difference between analogies and metaphors is not theoretical but rather technical and basically depending on the number and quality of mappings between the target and source domain. Adopting this perspective, we understand analogies and metaphors as comparisons that construct a similarity between two objects and we do not distinguish between those two notions. This definition will allow us to treat the RSA in the broader framework of metaphor analysis. More generally, understanding the nature of analogy and metaphor is a process central to scientific models and modelling (Gilbert, 2004). For the purpose of this study, we refer to models as artifacts which may be interacted with or visualized and we treat analogies and metaphors as one particular form of model in science education.

The increased interest in metaphors and analogies in science education stems partly from the fact that these models play an important role in scientific knowledge construction. There is a long tradition in the philosophy of science to argue for the epistemological importance of analogies (Hesse, 1953). Kapon and diSessa noted that "the generation of analogies and the reasoning stemming from these analogies play a central role in scientific practice, thought, and creativity" (2012, p. 262). Stinner (2003, p. 340) observed that the big theories in science including Einstein's theory of relativity or Maxwell's theory of electromagnetism are often the product of imaginative thinking which, according to Stinner, includes "to see analogies between disparate events". Thus, historical accounts of scientific discoveries abound with examples of how scientists used metaphors and analogies to build their theories (Chandler, 1994; Hesse, 1952; Kind & Kind, 2007; Silva, 2007). It seems that Einstein was particularly apt at finding fruitful analogies. He was presumably the one to introduce the RSA to reason about curved spacetime (Hentschel, 1998), and he used the analogy of riding on a ray of light to work out his theory of relativity in the first place (Kind & Kind, 2007).

Systematic metaphor analysis (Schmitt, 2005) is a recent fruitful approach that draws on findings from cognitive science and linguistics to understand the use of analogies and metaphors in science education (Amin, Jeppsson, & Haglund, 2015; Lancor, 2014a; Niebert & Gropengießer, 2014; Niebert et al., 2012). This approach goes back to Lakoff and Johnson who, in their seminal work (2003), argued that metaphors are not only a linguistic phenomenon, but a fundamental feature of thought and mind. Forming the basis of our conceptual systems, metaphors serve as a principal vehicle for understanding, because we systematically use inference patterns from one conceptual domain to reason about another conceptual domain. Since such metaphors are grounded in the everyday human experiences of "having a physical body in a physical world" (Roth & Lawless, 2002, p. 336) Lakoff and Johnson suggested that cognition is ultimately embodied. Embodied cognition extends the boundaries of the mind from merely being inside the brain to including the body's physical interactions with the world. Metaphors are thus the mediators that extend one physical experience to other conceptual domains. For example, the 'leg' of a table is an extension of the leg of a body, and allows us to make sense of its function as a structure for support (Lakoff & Johnson, 2003, p.??). We think about table legs in terms of our bodily experiences of being supported by our own legs and feet. Metaphors are thus not merely comparisons between two different things or concepts, but are rather frames through which we perceive and make meaning of the world (Schön, 1979). Applying systematic metaphor analysis through a perspective of embodied cognition highlights the bodily and experiential aspects of metaphor use.

The position that knowledge is embodied and that metaphors can reveal fundamental conceptions allows science educators to study learning processes through the lens of embodied cognition. Amin et al. (2015) acknowledged the emergence of a critical mass of studies that apply ideas from the perspective of embodied cognition in science education. These applications entail investigations into how the use of language and gestures can support conceptualization of abstract scientific ideas. We want to draw on those findings and employ similar methods to investigate the metaphorical patterns of the RSA in order to figure out in which ways students map basic features of the rubber sheet metaphorically onto the abstract scientific concept of spacetime. Embodied cognition does not imply that bodily understanding in some way supersedes the role of language in cognition, but rather suggests that language use and bodily understanding are intertwined.

Exploring the conceptual domain of GR from a linguistic perspective resonates with a broader movement in science education that emphasizes "talking science" in the classroom (Lemke, 1990). Reaching ultimately back to Vygotsky (1962), the assumption that language and the development of abstract thoughts are interrelated has brought about fruitful approaches to scaffold learners' development of scientific knowledge (Chen, Park, & Hand, 2016). Viewing language as a "window in the conceptions of students" (Niebert & Gropengießer, 2014, p. 281) aligns particularly well with the objective of our study: students are not familiar with the mathematical language of GR and have to reason by using the everyday language available to them to talk about abstract relativistic concepts. Metaphors are one particular example of talking physics. By choosing metaphors as our unit of analysis, we are able to employ a powerful linguistic tool to explore students' conception in GR.

2.2. Metaphor Analysis as an Analytic Framework in Science Education

An important study to employ embodied cognition as a framework in science education investigated students' struggles to understand analogies and metaphors as intended by teachers and instructors (Niebert et al., 2012), by reanalyzing 199 instructional analogies and metaphors on the basis of a metaphor analysis. By recognizing metaphors as a useful part of the material that can be analyzed and integrated into a broader research strategy, Schmitt (2005) proposed a systematic procedure for the reconstruction of metaphors to uncover patterns of thought. Niebert et al. built on this procedure to identify and classify conceptual metaphors in science education by first grouping metaphorical terms with the same source and target area and then summarizing the metaphorical model on the level "target is source". For instance "the gene is a code" and "equilibrium is a dance" are popular metaphors in biology and chemistry textbooks (Niebert et al., 2012). Their findings suggest that good analogies and metaphors in science education need embodied sources. This conclusion is an interesting one in light of the observation that the embodied source of the RSA and students' embodied understanding of gravity confront students with profound imaginative challenges; the analogy prompts students to transfer embodied understandings of gravity between 3 and 4 dimensions (Steier & Kersting, n.d.).

Research studies applied the concept of conceptual metaphor in a variety of ways, and developing a specific and operationalized definition of conceptual metaphor is a challenging but necessary task (Treagust & Duit, 2015). In the context of science education, Niebert et al. defined a conceptual metaphor as the "imaginative principles behind the analogy or metaphor" (2012, p. 855) that becomes apparent once metaphors and analogies have been arranged according to their target and source domain. That is, conceptual metaphors allow learners to imagine one thing in terms of another. Likewise, Lancor (2014a, 2014b) understood a conceptual metaphor as an overarching relationship between target and source domain that is supported by explicit metaphors/analogies that highlight or obscure characteristics of the scientific concept.

Metaphors and imagination are closely linked because metaphors mediate imaginative processes. Approaches to imagining depend on the notion of presence. As Nemirovsky, Kelton, and Rhodehamel defined it (2012, p. 131), imagining is the "experience of bringing to presence something which is absent in the current surroundings of the participants (Casey, 1979; Sartre, 2004)". Imaginers are interacting with objects, ideas, and situations that are not immediately *there* or perceivable. Metaphors, then, function as a way to give presence to these objects of imagination. Niebert et al. explained: "we employ conceptions from a source domain (…) and map them onto an abstract target domain (…) to understand abstract phenomena. Thus, the use of imagination requires a source–target mapping" (2012, p. 852). This imaginative mapping occurs through metaphor. One example used by Niebert et al., (2012) is the metaphor that *atoms are solar systems*. The abstract, difficult to visualize properties of an atom may become present for learners by relating atoms to the more concrete or familiar models of the solar system. We may imagine an atom (including its difficult to perceive properties), through metaphor, by drawing on our previous experiences with physical models of the solar system. By analyzing the structural properties and relationships of metaphor use, we are thus able to gain insight into how learners conceptualize, imagine, and make present abstract ideas.

While Niebert et al (2012) presented a broad picture of understanding instructional analogies in science education, other studies have used systematic metaphor analysis to focus on metaphors for individual scientific concepts such as the greenhouse effect or energy. Niebert and Gropengießer (2014) employed metaphor and qualitative content analysis to gain insight into students' and climate scientists' resources for understanding the greenhouse effect. Lancor (2014a) studied conceptions of energy in biology, chemistry, and physics and demonstrated that metaphor analysis can be a fruitful framework to analyze scientific discourse. She took a closer look at the substance metaphor for energy in textbooks and the science education literature and identified six conceptual metaphors within this broad metaphor: "Energy as a substance that can be accounted for, can flow, can be carried, can change forms, can be lost, and can be an ingredient, a product or stored in some way." (Lancor, 2014a, p. 1245) This analysis in turn helps to investigate how students understand science content, since each conceptual metaphor affords a different understanding of a scientific concept.

Since both the greenhouse effect and energy are particularly abstract concepts in science education, the above studies suggest that abstract scientific concepts might be too complex to be described by just one metaphor or analogy. Rather, they seem to be embedded in a metaphorical network that structures our understanding of a scientific concept (Lancor, 2014a); this is an observation that mirrors Lemke's (1990) suggestion that scientific concepts do not exist as ideas in their own separate reality, but that they are thematic items that make up a semantic pattern of relationships of meaning. This observation encourages us further to employ the framework of conceptual metaphors and embodied cognition in our study of the abstract concept of curved spacetime.

2.3. The Bad Use of Metaphors and the Use of Bad Metaphors

Ultimately, studying the role of metaphors in science education has the goal to improve instructional practices. In a recent editorial in this journal, Kampourakis (2016) pointed out that science educators have an important contribution to make: in communicating scientific knowledge, they bridge the gap between experts and non-experts. The use of metaphors plays a crucial role in this translation process. Calling for an increased awareness for the inherent limitations of metaphorical language and for the pitfalls that come with communicating conceptual issues, Kampourakis invited us to study "the bad use of metaphors and the use of bad metaphors". Genes are one example of the "bad use" of metaphors in biology education. According to Kampourakis, the popular metaphors of information encoded in DNA and the genome as a book of life can be misleading: those metaphors present genes as autonomous entities without taking the cellular context into account. One has to be explicit in communicating that encoding information is not an inherent property of genes.

Kampourakis' call created a common interest in metaphorical practices to which we aim to contribute with this study. Investigating how the – possibly "bad" – RSA can be put to good use in teaching and learning of GR is very much in line with a recent exploration by Haglund (2017), who studied the scientific concept of entropy that is metaphorically conceptualized as disorder. Just like spacetime, entropy is "a genuinely challenging concept for students to grasp, due to its abstract, complex, and mathematical nature" (Haglund, 2017, p. 208). Haglund argued that the disorder metaphor can give a first flavor of entropy that students in turn can use to develop and refine their understanding of entropy.

In contrast to entropy, the notion of curved spacetime, although abstract and mathematical in nature, is intimately linked to the embodied experience of being under the influence of gravity. Coming to full circle with the starting point of our investigation, we wish to understand how learners conceptualize their experience of gravity in the setting of GR.

3. Methods

Before we can explore the ways in which students conceptualize gravity and curved spacetime with the help of the RSA, it is important to have a sound understanding of the RSA. Therefore, our methodological approach entails the analysis of two different data sets: first, we use metaphor theory to analyze the rubber sheet analogy based on the general accounts of physicists and physics educators as found in the literature. These findings serve as basis for the second part: our empirical investigation of students' use and understanding of the RSA in relation to gravity and curved spacetime.

3.1. Metaphor analysis in RSA-relevant literature

To study the presentation of the RSA in the relevant literature, we followed the systematic procedure for the reconstruction of metaphors as outlined in Schmitt (2005) and further refined in Niebert et al. (2012). This approach promotes the analysis of metaphors to a qualitative research procedure that allowed us to reconstruct metaphorical concepts based on written accounts.

The two crucial steps in a systematic metaphor analysis consist in: 1) identifying a metaphor and 2) reconstructing metaphorical models (Schmitt, 2005). First, to identify metaphors, one looks for phrases that can be understood beyond their literal meaning, which stems from physical or cultural experience (source area) and is transferred to a new, and often abstract, area (target area). Second, to reconstruct metaphorical models, a process that Niebert et al. (2012) called “categorizing the level of conceptual metaphor”, one groups the metaphorical phrases that have the same source and the same target area. Condensing this categorization in the equation “target area = source area”, one thus reconstructs the complete metaphor by identifying its underlying logic.

To exemplify the process of the systematic metaphor analysis, we look at an exposition from Baldy (2007, p. 1772) that makes the mapping between target and source area in general relativity very specific:

> Einstein’s theory is introduced to students via the so-called “pillow” model: the pillow represents space, and steel balls of different sizes and masses are used to represent celestial bodies. When a marble representing a body is placed next to a ball, it falls into the dip in the pillow created by the ball. And if the marble is rolled fast enough, it deviates from its normal trajectory in the vicinity of the ball.

Here, the identification of metaphors reveals a rich network of source and target areas that, furthermore, interact dynamically. We can identify several source areas rooted in everyday experience – namely a pillow, a dip in the pillow, a steel ball, and marbles. We find three abstract target areas – space, celestial bodies, and trajectories. To structure the analogy on the level “target-is-source”, we can formulate “space is pillow”, “steel balls are celestial bodies”, and “marbles are celestial bodies”. In addition to these mappings of objects, we have another dimension to the metaphor, namely the dynamic interplay between target and source objects: a ball creates a dip in the pillow, a marble falls into the dip, a marble deviates from its trajectory. We return to this example in our presentation of the results in the next section.

Since science educators are not only interested in identifying analogies and metaphors in scientific discourse, but are also concerned about communicating scientific ideas fruitfully, one can extend the systematic metaphor analysis in a way that encompasses educational concerns. Niebert et al. (2012) proposed two additional steps as part of an extended metaphor analysis that is valuable in the educational context: the identification of the metaphor’s deficiencies and resources, and the comparison and interpretation of students’ and teachers’ source domains. We incorporate these two steps in our analysis, noting that they allow us to make the transition from our literature review to the empirical interpretation of students’ conceptual understanding.

Since we conducted this study in the context of the Norwegian physics curriculum, our selection of relevant texts for a metaphor analysis of curved spacetime includes the two Norwegian physics textbooks on the market (Callin, Pålsgård, Stadsnes, & Tellefsen, 2012; Jerstad et al., 2014), two popular science books by renowned physicists in the field of general relativity (Greene, 2010; Thorne, 2009), six peer-reviewed research articles that address the RSA explicitly and that were published within the last 25 years (Baldy, 2007; Chandler, 1994; Gould, 2016; Kaur et al., 2017a; Middleton & Weller, 2016; Price, 2016), as well as one master’s thesis in science education (Watkins, 2014). Following the systematic procedure as outlined above, we identified 41 instances of metaphorical phrases that relate the scientific concept of curved spacetime to a rubber sheet-like object. To simplify the classification in terms of “target-is-source”, we further structured the metaphorical phrases with the help of three subcategories “spacetime is”, “objects are”, “dynamical action via”. This subdivision follows our observation in the previous example that there is a metaphorical mapping of target and source objects, as well as a dynamical interplay between the two. Based on this subdivision, we were able to identify four conceptual metaphors for curved spacetime that allow for a full reconstruction of the RSA. After having unpacked the presentation of the RSA in this way, we followed the extension of the metaphor analysis by Niebert et al (2012) in order to identify deficiencies and resources of each conceptual metaphor: we took into account the strengths and weaknesses of the RSA that were mentioned explicitly in the analyzed literature and compared those to the individual conceptual metaphors that make up the RSA in order to identify features that the RSA possibly highlights or hides.

3.2. Metaphor and thematic analysis of students’ responses

With the systematic metaphor analysis of the literature, we have laid the groundwork for investigating students’ conceptualization of gravity and curved spacetime. Before we explain how the literature analysis has informed the way that we framed the empirical analysis, we outline the data collection procedure and the greater educational research project that this study is part of.

3.2.1. Data collection

This work was conducted within the design-based research project ReleQuant that developed collaborative online learning environments in modern physics for upper secondary schools in Norway (Henriksen et al., 2014). Drawing on the tradition of Vygotsky (1962), project ReleQuant builds on a sociocultural approach to learning physics that emphasizes the use of language (Lemke, 1990; Scott & Mortimer, 2005) and the interdependence between the individual student and his or her surroundings in the learning process (Rasmussen & Ludvigsen, 2010). Students were encouraged to work in pairs or small groups and discuss key concepts of GR and quantum physics while using the learning environments.

This study reports findings from the second round of testing a GR learning environment in five upper secondary physics classes in three Norwegian schools that were considered to be high achieving in national comparison. The schools are partner schools of the ReleQuant project and the teachers were involved in the development of the learning resources that were jointly designed by physics educators and learning scientists from the project. In total, 97 students (70 boys, 27 girls, 18-19 years old) participated in a series of two 2-hour lessons that were part of the regular physics curriculum for final year secondary school students in Norway. The curriculum states that students should be able to “give a qualitative description of general relativity” (The Norwegian Directorate for Education and Training, 2006). The learning environment consists of three thematic units the last of which covers the topic of curved spacetime.

Our interest in understanding student ideas of the RSA led us to choose one particular discussion task, which addresses the RSA directly, for further analysis (figure 2). The task invited students to reflect on the RSA by discussing a cartoon that addresses the analogical nature of the rubber sheet representation. The open format of the question is well suited to investigate students’ ideas of curved spacetime and prompts them to consider the role of analogies more generally. In a second step, students had to write a short summary of their group discussion. This summary provides insight into their use of scientific language, as well as what they felt were the most important conclusions in their discussions.. Our data comes from 65 written responses to this retrieved from the online learning platform. The reason that the total number of collected written responses (65) is smaller than the total number of participating students (97) is that several groups of students chose to submit a joint group response instead of writing individual summaries.

The discussion task is part of a longer learning sequence that introduces students to the concept of spacetime by presenting different models and interactive visualizations of curved spacetime. In the discussion, we relate the findings of this study to the broader context of investigating students’ conceptual understanding of spacetime.

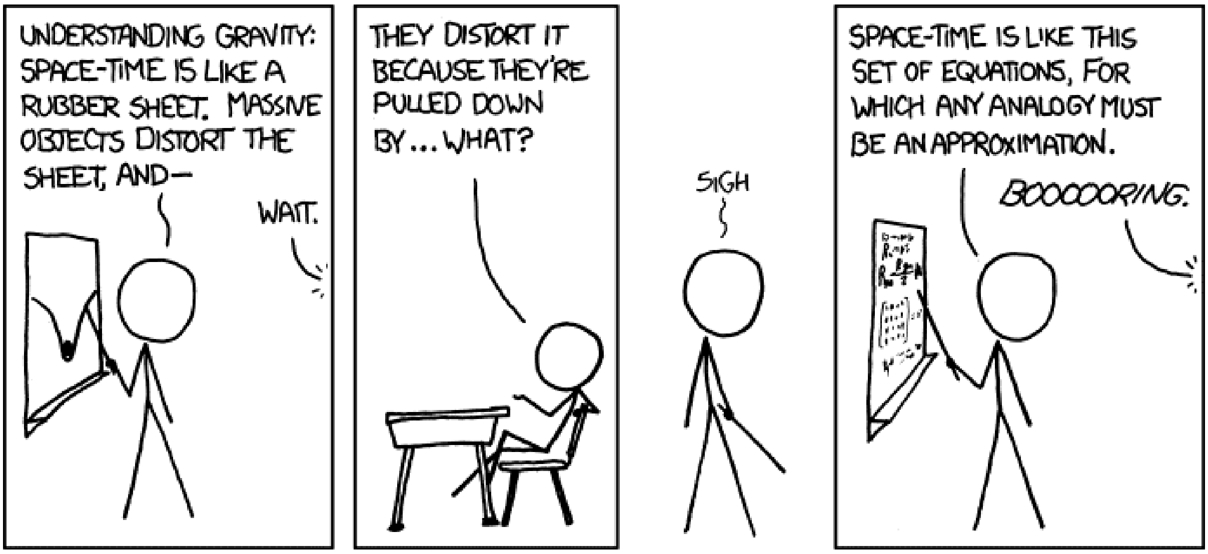


Fig. 2 In the spacetime unit, students worked on the following task: "Use your knowledge in general relativity and discuss the cartoon. Write down a short summary of what you have discussed." The comic is licensed under a Creative Commons Attribution-NonCommercial 2.5 License and can be accessed under https://xkcd.com/895/

3.2.2. Data analysis

We conducted two independent analyses of students' responses: a metaphor analysis to identify conceptual metaphors and a thematic analysis to characterize students' awareness of the strengths and weaknesses of the RSA. This double approach resembles the one employed by Lancor (2014b), who characterized students' conceptual understanding of the energy concept through the lens of metaphor analysis.

Employing methods of the systematic metaphor analysis and mirroring our procedure of the systematic metaphor analysis of the literature, we took students' written responses and identified 39 instances of metaphorical language connected to curved spacetime. The metaphorical phrases were thus again divided into the three subcategories "spacetime is", "objects are", "dynamical action via". Following the scheme "target-is-source", we then continued to decompose each metaphorical phrase into its various mappings between target and source area.

We found students' responses to be often somewhat muddled and not very clear about mappings between target and source areas. To deal with this ambiguity in the written responses, we were very careful in conducting the metaphor analysis. In particular, we found many phrases in which students used a kind of rubber sheet analogy without directly mapping from the target to the source area, i.e. they remained either in the target or in the source area. Therefore, we chose to generate an additional code for implicit metaphorical mapping and tagged 22 instances of those phrases in addition.

To illustrate our method, we present an example of analyzing two responses:

> Student response 1: If you put a mass on a sheet it will bend and create a deflected/curved spacetime around the mass like that we have seen before where the sheet was time.

> Student response 2: It has to do with that the mass of an object went down in the paper as it was described.

In response 1, we can identify the mappings "sheet is spacetime", "sheet is time", "spacetime bends", "spacetime is deflected/curved". Response 2 is an example of an implicit mapping, because the student remains in the source area of paper and mass without explicitly mentioning spacetime or celestial objects. Nonetheless, we can identify the conception "mass goes down in the paper".

Since we were not only interested in the way students' conceptualize curved spacetime linguistically, but also wanted to gain insight into their awareness of the analogical nature of the RSA as well, we conducted an additional thematic analysis (Braun & Clarke, 2006) of the data set to unpack students' understanding thereof.

This analysis corresponds to the additional step of identifying strengths and weaknesses of metaphorical mappings as suggested by Niebert and Grobengießer (2014). However, the important difference to the corresponding analysis of the literature is that we aimed to bring to light students' *own* ideas of strengths and weaknesses of their metaphorical reasoning, instead of reconstructing general features that the RSA highlights and hides.

Following the five step procedure of a thematic analysis (Braun & Clarke, 2006), we got familiar with the data set by coding it for general occurrences of students' elaboration on analogies or scientific models. Based on our literature analysis, we used the identified strengths and weaknesses of the RSA as starting point to generate a set of initial codes. With this set we analyzed the identified analogical responses and started to create new codes that captured recurring patterns in the responses that we could not have anticipated solely from the metaphor analysis of the literature. For example, eight groups of students addressed the interaction between the student and the teacher in the cartoon. This observation gave rise to the code "teaching situation". With this enriched set of codes, it became evident that we could group student responses dealing with the RSA into three themes: responses that elaborate on the general nature of analogies in physics, those that address specific characteristics of the RSA, and those that comment on the context of the cartoon. Based on this broad classification, we reviewed and refined our codes and coded the data set again. The final set of themes and codes is presented in figure 3.

The first author conducted the first two steps of the thematic analysis and identified the relevant responses for the metaphor analysis. To ensure the validity of the analysis, both authors then discussed the mappings and the codes over several rounds while reviewing all responses together until they reached agreement. Particular focus was put on the interpretation of the findings that were critically re-examined in light of the literature findings.

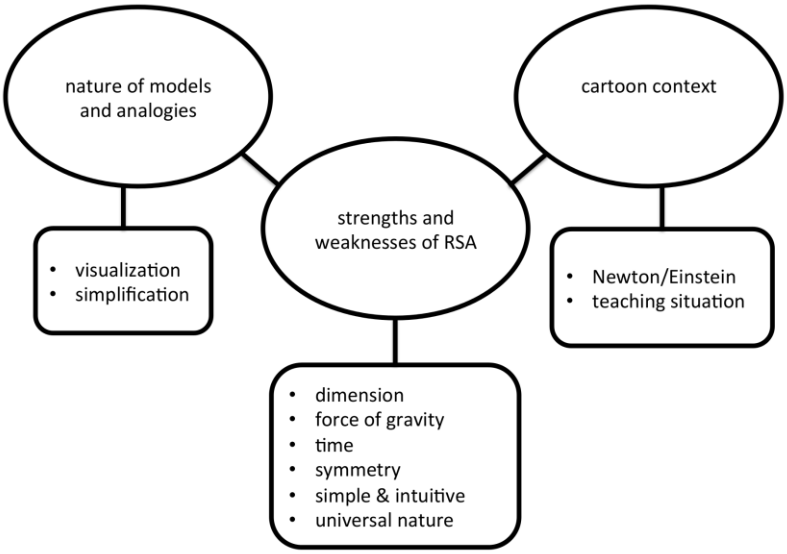


Fig. 3 Map of themes and codes of the thematic analysis of student responses

4. Results

In this section, we attempt to spell out the nature of the RSA and characterize upper secondary students' understanding of it based on a metaphor analysis of relevant literature and a combined metaphor and thematic analysis of students' written accounts. Following this dual approach, we present results from the literature analysis first and use these results to contextualize the empirical findings from students' responses.

4.1. Metaphor Analysis of the RSA According to the Literature

The goal of our systematic metaphor analysis was to structure the RSA on the level "target-is-source" and to identify and reconstruct the conceptual metaphors that guide this classification. Based on our analysis of relevant literature, we were able to unpack the metaphorical network of the RSA by identifying four different conceptual metaphors. Each of these conceptual metaphors affords understanding of a different aspect of the concept of gravity as curved spacetime by highlighting and hiding various features of the scientific concept. In table 1, we give an overview of the systematic metaphor analysis of the literature.

Table 1 - The RSA encompasses four conceptual metaphors each of which can be exemplified by specific analogies. The conceptual metaphors are synthesized from a systematic metaphor analysis of relevant literature. The examples come from analogies found in the literature. The conceptual metaphors that comprise the dynamical mapping can be formulated either from the spacetime or the mass perspective.

Conceptual metaphor			Analogies that exemplify the conceptual metaphor
Static mapping	Spacetime is a **fabric** that is **malleable**.		Spacetime is a **piece of rubber** that is distorted. Spacetime is a **trampoline** that is stretched. Spacetime is a **pillow** that is deformed. Spacetime is a **membrane** that is warped.
	Spacetime is a **2D-surface** that has **geometrical features**.		Spacetime is **flat.** Spacetime is **curved**. Spacetime is **bumpy.** Spacetime has **slopes.**
Dynamical mapping	Spacetime is a **background** that **responds** to the presence of massive objects.	Massive objects **distort** spacetime.	Spacetime **stretches down** under the weight of an object. Spacetime **bends in** towards an object. Objects **create cavities, slopes and depression** in spacetime.
	Spacetime is an **actor** that **influences** the movement of objects.	Objects **move under the influence** of spacetime.	Objects **roll across** spacetime. Objects **fall in towards** heavy objects. Spacetime curvature **alters the path** of objects. Objects **deviate from their trajectory** in response to deformation.

Before looking closer at the four conceptual metaphors that the RSA encapsulates, we want to make two preliminary remarks. First, it is important to note that there are mappings in the RSA that seem to be less interesting with respect to the characterization of gravity as the geometry of spacetime. While we have identified many examples of objects that are commonly placed on the rubber sheet such as bowling balls, golf balls, marbles, and rocks, these objects do not reflect a relevant imaginative principle that characterizes one thing in terms of another, but are just examples of massive objects that exert a gravitational effect. Even though we might say that the *bowling ball* curving the rubber sheet is like the *sun* curving spacetime, this comparison is mostly an upscaling from everyday size objects to cosmic scale objects. However, the intrinsic feature of being a massive objective does not change when going from a ball to the sun. Thus, the mapping is qualitatively different from the mapping that takes place on the level spacetime-is-rubber sheet. When identifying conceptual metaphors for gravity, we therefore focused on the target-is-source mappings that deal with spacetime itself.

Second, as noted already in the methods section, the RSA entails two different kinds of mappings. First, there is a static mapping that maps an experience-based source area like the rubber sheet and marbles to the target area of spacetime and planets. Second, there is a dynamical mapping that encodes the dynamic interplay between the different actors of the mapping, i.e. how masses curve spacetime just like marbles curve a rubber sheet. It is this dynamical interplay that gives rise to the phenomenon of gravity. We argue that both types of mappings are important and constitute a metaphorical network of gravity as curved spacetime. The static mapping settles the underlying structure of the RSA, whereas the dynamical mapping employs the "basic logic" (Niebert & Gropengießer, 2014, p. 299) of the source domain to make sense of the physical mechanism of gravity.

To exemplify the four identified conceptual metaphors below, we will use the following example from a physics education article:

> (…) let us now think of spacetime as though it were a rubber sheet stretched on a frame hanging over the ground. If there is no matter in it, spacetime is flat. If a particle, a marble or a light ray, were rolled across flat space time, it would go straight. If, on the other hand, matter, a star for example, is present, it acts like a weight on the sheet and creates a distortion. The sheet would stretch down under the weight; the greater the weight the greater the indentation. Now when a marble or light ray is rolled across the sheet it curves into the depression. In this picture the particle is moving rapidly enough to bend in toward the lump and continue to move on out. Another particle might circle and eventually fall into the depression. (Chandler, 1994, p. 171)

While spacetime as the target domain remains the same, we have found a variety of source domains that get mapped onto this abstract domain: a rubber sheet, a pillow, a membrane, a trampoline. However, all source domains have one feature in common which leads to the reconstruction of the first conceptual metaphor: **Spacetime is a fabric that can be stretched and deformed.** This conceptual metaphor captures the idea that all source domains are fabric-like objects that are malleable. Evidence for this conceptual metaphor includes the use of a source domain that either implicitly displays this property (as for example a rubber sheet does) or explicitly mentions the stretching and deforming of the source domain: "(…) let us now think of spacetime as though it were a **rubber sheet stretched** on a frame hanging over the ground."; "The sheet would **stretch down** under the weight; the greater the weight the greater the indentation."

Moreover, most mappings did not stop at the level of comparing spacetime to a fabric. The internal logic of this mapping invites us to deduce further characteristics of the target domain, which leads to the formulation of the second conceptual metaphor: **Spacetime is a two-dimensional surface that has geometrical features.** In the literature, we found analogies that characterized spacetime via a source object that is flat, curved, bumpy, twisted, has a slope, and which, accordingly, has geometrical features. These characterizations imply in particular that spacetime is a two-dimensional surface embedded in three-dimensional space: "If there is no matter in it, spacetime is **flat**. If a particle, a marble or a light ray, were rolled across flat space time, it would go straight."; "In this picture the particle is moving rapidly enough to bend in toward the **lump** and continue to move on out."

These two conceptual metaphors make up what we call the static mappings of the RSA. They characterize spacetime in terms of more familiar notions, but do not yet explain how gravity arises. The explanation of this phenomenon is captured by two additional conceptual metaphors that make up the dynamical part of the mapping. Note that each of these two conceptual metaphors can be formulated either from the spacetime or the mass perspective: **Spacetime is a background that responds to the presence of massive objects/Massive objects distort spacetime.** In the literature, the RSA is used to explain how gravity arises by saying that spacetime stretches down under the weight of objects, spacetime bends in towards objects, or that it is distorted by objects. On the other hand, it is said that objects create cavities, slopes, or depressions. Mappings were considered to have evidence of this conceptual metaphor if they discussed either the way that spacetime reacts to the presence of massive objects or the distortion effect of massive objects on spacetime: "If, on the other hand, matter, a star for example, is present, it **acts like a weight** on the sheet and **creates a distortion**. The sheet would **stretch down under the weight**; the greater the weight the greater the indentation."

Finally, we have the metaphor: **Spacetime is an actor that influences the movement of objects/Objects move under the influence of spacetime.** Evidence for this conceptual metaphor entails the way objects react to the geometry of spacetime or the way curvature alters their paths: objects deviate from their trajectories, they curve or fall in towards massive objects, and their motion changes in response to deformation: "Now when a marble or light ray is rolled across the sheet **it curves into the depression**. In this picture the particle is moving rapidly enough **to bend in toward the lump** and continue to move on out. Another particle **might circle and eventually fall into the depression**."

By definition, conceptual metaphors capture the underlying relationships that guide analogical mappings between the target and source domain. Thus, breaking down the ways that gravity is conceptualized in the RSA helps to identify the strengths and weaknesses of the analogy. In order to do so, we compared the strengths and weaknesses of the four conceptual metaphors that we had identified that were mentioned explicitly in the literature. This comparison allowed us to supplement the literature collection of strengths and weaknesses with our own findings. To answer our first research question, we synthesized the features that the RSA brings into focus and obscures in table 2.

Table 2 - Strengths and weaknesses of the RSA. Findings in the table were synthesized from literature examples supplemented with our own findings based on a metaphor analysis of the literature.

Strengths of the RSA	**Weaknesses of the RSA**
• spacetime is dynamic and not static • spacetime is influenced by objects • spacetime alters the movement of objects • gravity exhibits a universal nature: spacetime responds universally according to the weight of the objects - the more massive an object, the more distortion of spacetime it will create • gravitational phenomena involve no "mysterious" action at a distance • gravity is geometry; the RSA provides a mechanism of how gravity arises • the RSA is a simple, intuitive model • the RSA has great explanatory power; it is suitable to show orbital motions, curved space, and photon trajectories	• the RSA obscures that spacetime is four-dimensional • the RSA obscures that spacetime has a temporal dimension • the RSA obscures that curvature is an intrinsic feature; it depicts curved spacetime as if there was an unseen dimension into which spacetime curves • the RSA obscures that curvature around massive objects is symmetric in all dimensions • the RSA makes use of the force of gravity to explain the distortion of the rubber sheet

4.2. Students' Understanding of the Rubber Sheet Analogy

We found a big variety in students' written responses in terms of length, depth of reflection, and the range of issues addressed. This variety shows that students engaged with the task in many different ways. The task was an open one: by asking students to use their knowledge of GR to discuss the cartoon and to summarize their discussion in written form afterwards, we challenged them to figure out what they felt was important. In 39 of 65 responses, we identified instances of metaphorical language that were accessible to metaphor analysis, whereas 42 of 65 responses addressed limitations and strengths of the analogies. Those responses encompassed elaborations on the need to employ analogical reasoning in science, as well as pointed out specific shortcomings of the RSA in the context of GR. In addition, 12 responses dealt with the instructional context of the cartoon and how the interaction between teacher and student contributed to understanding GR.

It was interesting to see how students incorporated different parts of the learning environment in order to solve the task. Many connected the cartoon to explanations previously presented, such as our inability to visualize four dimensions except mathematically. Thus, the format of the question seems to have been successful in engaging students to piece together the different bits of explanations that convey the complex scientific concept of gravity as curved spacetime.

4.2.1. Systematic Metaphor Analysis of Student Responses

To characterize students' understanding of the RSA, we first looked at the ways in which students talked about the RSA. Analyzing the language they employed through the lens of metaphor analysis allowed us to approach our second research question.

In the metaphor analysis of the literature we found four conceptual metaphors that describe the relationships between spacetime and massive objects. Conducting a similar analysis of students' responses, we found that students displayed a wider range of target-source-mappings (table 3, figure 4). While the literature only identified productive target source mappings, students had not acquired a complete understanding of the analogy yet, and were therefore likely to produce mismatches between target and domain areas of the analogy. Naturally, students produced more mappings because there are many possibilities to create mappings between target and source objects. However, on a deeper level, these mismatches allowed us insight into the challenges that students face when conceptualizing gravity and spacetime.

Similarly to the characterization of the literature findings, we divided the student-generated mappings into static and dynamic ones (table 3). In general, occurrences of static mappings were less frequent than dynamical ones and there was a greater variety of mappings in the dynamical domain. This difference in frequency provides a first hint that students displayed more misconceptions in the dynamical mappings of the RSA. They might struggle most with the actual mechanism of gravity (i.e. the dynamical interplay between target and source components that give rise to the physical phenomenon of gravity) than with the static mapping between spacetime and rubber sheet as such.

Most of the static mappings only broke down spacetime into space and time components. Students mapped both the space and the time component onto a fabric-like object that resembled a surface with geometric features. This object could be a rubber-sheet, a sheet, a trampoline, a tablecloth, a paper, or a rubber-pad; but no matter what actual source domain the students chose, their mappings resembled the two static conceptual metaphors we found in the literature. In one instance, students chose "lines" as the source domain to describe spacetime with. We interpret this choice as borrowing from a common way of depicting spacetime with the help of a deformed mesh (figure 1).

The only novel mapping we found that differed significantly from the common comparison of a rubber sheet to space, time, or spacetime involved the fourth dimension:

> Einstein thinks that objects with mass curve spacetime, the **fourth dimension**. He thinks that people live in a four-dimensional reality where **the fourth dimension is spacetime**. . (Student group 1)

This response shows how students confused the new terminology, which seems to be particularly challenging. Even though the formulation "mass curves spacetime" is a correct one, it becomes clear that this group of students still struggled with the abstract notions of the fourth dimension and spacetime both of which are equated in this response. Thus, using the right terminology could in some cases mask students' lack of conceptual understanding and a metaphor analysis allowed exploring whether this was indeed the case.

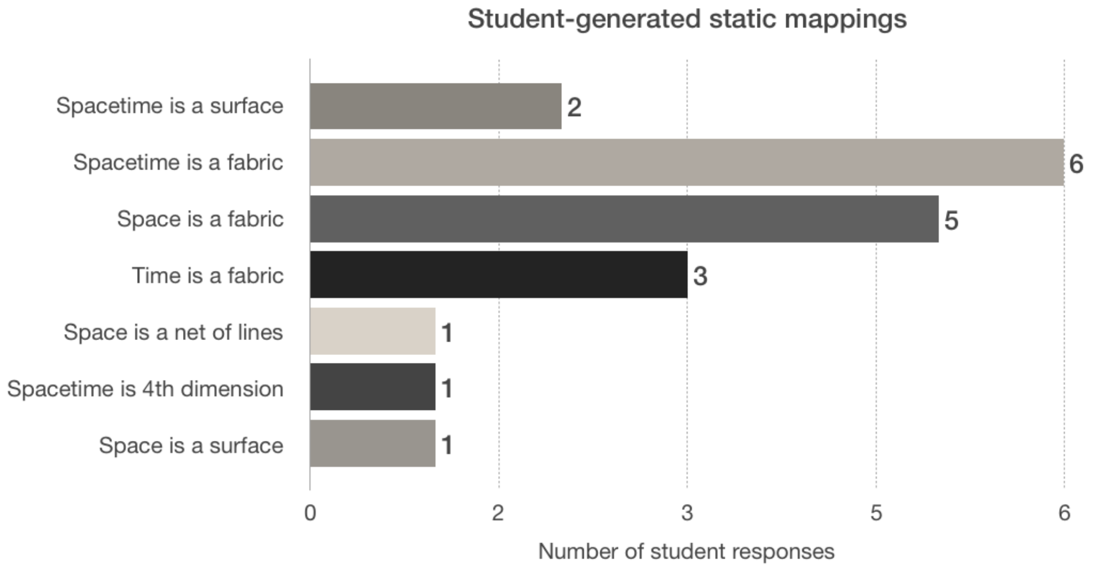


Fig. 4 Conceptual metaphors that students generated in the static mapping between source and target domain. The height of the bar represents the number of student responses. In total, 39 responses featured various metaphorical instances of which 19 corresponded to static mappings.

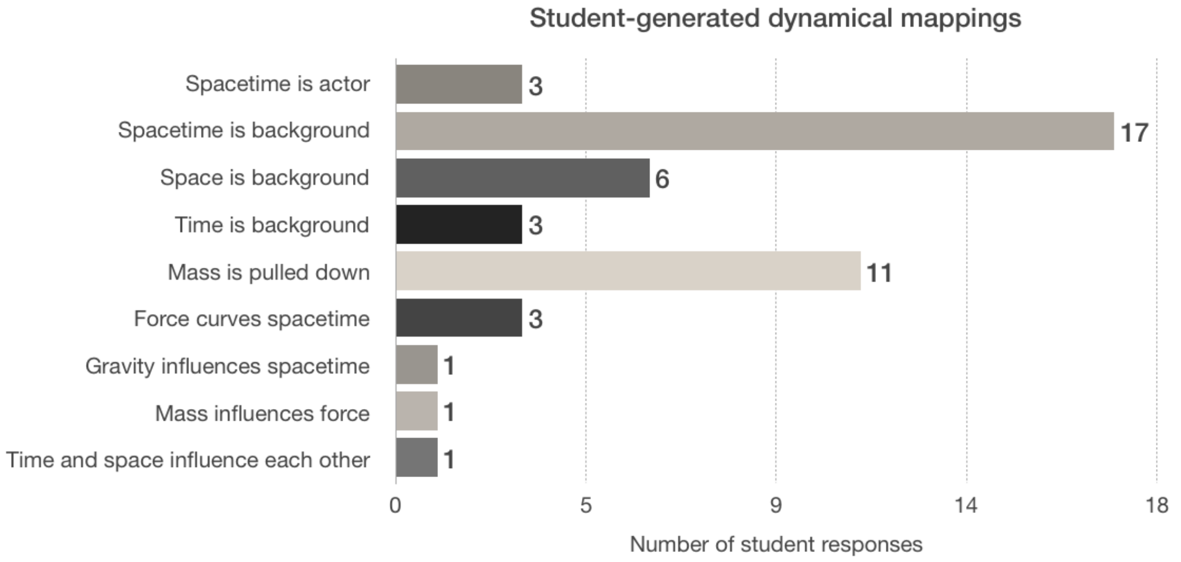


Figure 5 - Conceptual metaphors that students generated in the dynamical mapping between source and target domain. The height of the bar represents the number of student responses. In total, 39 responses featured various metaphorical instances of which 46 corresponded to dynamic mappings.

One would expect an added level of complexity when students have to describe the physics of gravitation that is captured by the dynamic relationships between target and source domains. Our findings align with this speculation, as students generated a greater variety of dynamical mappings (figure 5). While the most common mappings corresponded to the ones identified in the literature - namely that spacetime (or space or time separately) is a background that responds to the presence of massive objects and an actor that influences the movement of these - we identified various other novel conceptual metaphors. These conceptual metaphors concerned mainly the interplay of force, mass, and spacetime.

Not surprisingly, the most common novel conceptual metaphor addressed the problem that is featured in the cartoon: in order for the mapping to work, the RSA relies on the force of gravity that pulls a massive object down, thus explaining the relativistic notion of gravity with its classical counterpart. Of course, the task invited students to observe this. Accordingly, almost all responses that employed the passive (or Newtonian) perspective that the mass is pulled down into the sheet instead of the active (or Einsteinian) view that mass curves spacetime expressed criticism towards this idea. We come back to this observation in the next section when looking closer at students' awareness of the analogical nature of the RSA.

Less frequent but crucially related to the Newtonian conception of gravity is the idea that spacetime is curved by a force or by gravity acting on it:

> We discussed what this **force that curves spacetime** could be since it is not a force. *(*Student group 2*)*

This response and the general mappings that conflated forces with the analogical mappings show that students still used the force concept in their reasoning even though they "knew" and were told that gravity is not a force. These conceptual metaphors thus point towards a conceptual struggle that students faced when attempting the transition from classical to relativistic theories of gravitation. They confused cause and effect in the analogy: the force of gravity does not curve spacetime, but it arises from the curvature of spacetime.

Finally, we would like to comment on the implicit mappings that we already mentioned in the methods section. Many students used the RSA implicitly – 22 out of 39 metaphorical phrases remained either in the target or the source area. This observation could first of all be simply a sign of the fact that students inferred from the given context that the mapping was there without seeing the need to actually spell it out. But it could also indicate an insufficient understanding of what the target and the source domains were and might display lacking of mastery of the domain specific language. Possibly, the usefulness of analogical mapping was not clear to them - the productive use requires explanations of the relationship between target and source.

In table 3, we list all student-generated conceptual metaphors. Each metaphor is exemplified by a student response. It is important to note that student responses often comprise several conceptual metaphors, so our choice of examples does not necessarily reflect just one particular conceptual metaphor.

Table 3 - Student-generated mappings between target and source domains of the RSA. The shaded conceptual metaphors are the ones found in the literature. Examples are translations from student responses retrieved from the learning environment.

Conceptual metaphor		**Examples of student response.**
Static mappings	Spacetime is a fabric that is malleable.	In this cartoon we see that the teacher tries to explain spacetime by comparing it to a **rubber pad** where heavier masses fall further down than smaller masses. (Student group 19)
	Spacetime is a 2D-surface that has geometrical features.	Spacetime is influenced by gravity, therefore **the rubber sheet gets twisted.** (Student group 20)
	Space is a fabric that is malleable.	In the cartoon there is a question what pulls the object downwards such that the **space gets curved**. But you should not see this as a force, but that the **space "curves itself around"**. (Student group 21)
	Time is a fabric that is malleable.	If you put a mass on a sheet it will **bend** and create a **curved** spacetime around the mass like that we have seen before where **the sheet was time**. (Student group 22)
	Space is a surface that has geometrical features.	(…) space (…) can almost be viewed as **a sheet around the object**. (Student group 23)

	Space is a net of lines.	Here it is introduced that the “force” of gravity pulls the **lines** down. This is wrong according to Einstein. (Student group 24)
	Spacetime is the fourth dimension.	Einstein thinks that objects with mass curve spacetime, the **fourth dimension**. He thinks that people live in a four-dimensional reality where **the fourth dimension is spacetime**. (Student group 1)
Dynamical mappings	Spacetime is a **background** that **responds** to the presence of massive objects.	**Spacetime curves** itself around the masses because the masses “lie” on top of spacetime and **press it down**. (Student group 25)
	Spacetime is an **actor** that **influences** the movement of objects.	Mass curves spacetime and **spacetime determines therefore the movement** of the masses in spacetime. (Student group 26)
	Space is a background that responds to the presence of massive objects.	The point is not that the mass is “pulled down” in space. **Space curves itself around the mass**. (Student group 27)
	Time is a background that responds to the presence of massive objects.	Big **masses curve all of time** and space and do this in several dimensions. (Student group 9)
	Mass is pulled down.	This can be difficult to visualize, so we usually look at this in two dimensions. Then it looks as if there is **something that pulls the mass down**. But it is the mass itself that curves the space. (Student group 9)
	A force curves spacetime.	We discussed what this **force that curves spacetime** could be since it is not a force. (Student group 2)
	Gravity influences spacetime.	**Spacetime is influenced by gravity**, therefore the “rubber sheet” gets twisted. (Student group 20)
	Mass influences force.	The ball sinks down into the sheet because of gravity and heaviness, but in the outer space heaviness will not make it fall down. This is because there is no force of gravity in outer space. Instead, **the mass of an object will tell how much force of gravity** it has. How much it attracts other objects. (Student group 28)

4.2.2. Thematic Analysis of Student Responses

The metaphor analysis of students’ language served as a starting point from which we further explored students’ understanding of the RSA. The thematic analysis of student responses allowed us to move beyond the structural linguistic level by taking into account how students showed awareness for the analogical nature of the RSA. An overview of the frequency of codes is displayed in figure 6. In what follows, we explain the findings in detail.

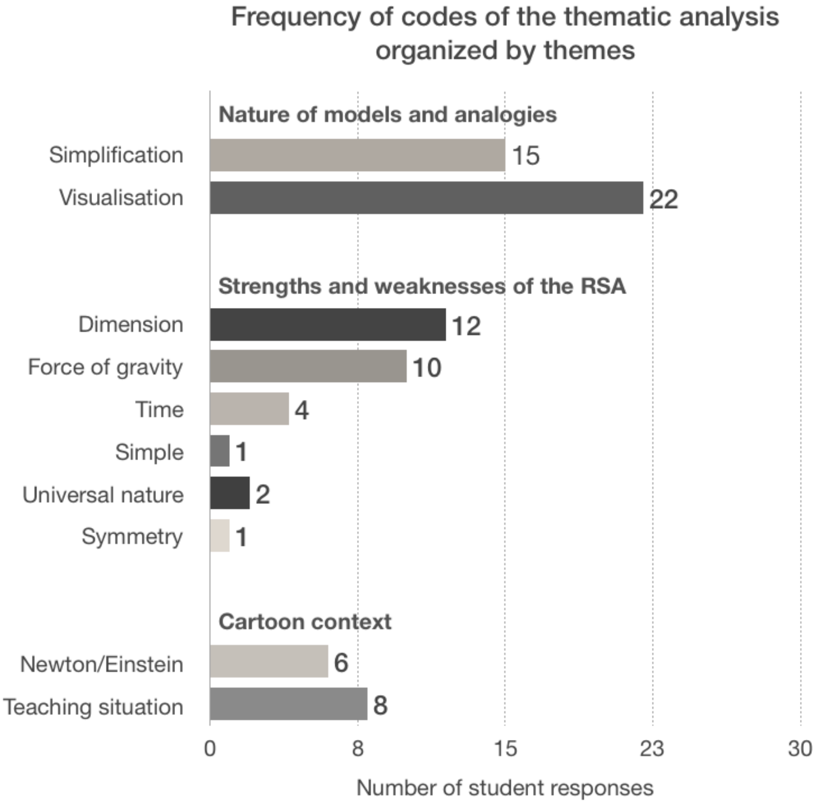


Fig. 6 The thematic coding of student responses comprises three themes and ten codes. In total, 42 responses featured student talk about analogies.

In the two most frequent types of responses, students displayed a general understanding for the role of analogies and analogical models in science, which we turned into the theme “nature of models and analogies” and that consists of the codes “visualization” and “simplification”. The code “simplification” encompasses written accounts that express the insight that analogies are always limited in their explanatory power and that they inevitably simplify or approximate a phenomenon to a certain extent:

> A useful tool to understand physical phenomena are models. The problem with the models is that they are simplifications. In this case the models become actually wrong. You could think that it is the force of gravity that pulls the object down, but there is no force of gravity. The alternative is to explain the phenomenon purely mathematically, but then you don’t have any illustration. *(*Student group 3*)*

Here, students displayed awareness of the limited nature of the RSA and expressed the understanding that models of gravity can only be an approximation, as well as that it is only through mathematics that one can fully describe GR. Other students were more explicit in relating the need for visualizations to their understanding of the simplifying function of models:

> We make models to describe physical phenomena, but these models are simplifications and not quite precise. They help us to visualize, even though they don’t tell the whole truth. Mathematically, we get the correct results just by using calculations, but to understand curvature of spacetime we need to visualize it with help of simplifications. *(*Student group 4)

In those two examples, we can also identify another important issue that got mentioned repeatedly: the inability to visualize curved spacetime. Students expressed their awareness for their inability to visualize more than three dimensions:

> *It is impossible to make a precise three-dimensional representation of a four-dimensional phenomenon.* (Student group 5)

> *It is impossible to visualize*[1] *four-dimensional spacetime, and you need to use two- and three-dimensional analogies that approximately can give an understanding of how four dimensions work.* (Student group 6)

> *We live in a four-dimensional world where three of them can be understood by human beings. To understand the concept of curvature of spacetime we can use analogies, but analogies will never make you visualize time, this can only describe the effect of spacetime curvature.* (Student group 7)

While students showed a quite sophisticated understanding of the need for analogies and visualizations in the domain of GR, many of the responses remained on a rather general level and only about half explained specific shortcomings of the RSA. These strengths and weaknesses that relate directly to the RSA are summarized in the second theme that encompasses six codes which we contrast with the significantly longer list of strengths and weaknesses as synthesized based on the literature analysis in table 2.

The most common limitation of the RSA that students identified was the reduction of a four-dimensional phenomenon to a lower-dimensional representation tagged by the code "dimension". While this weakness of the RSA is closely related to the general inability to visualize four dimensions, some students touched upon the problem of "intrinsic" curvature versus "extrinsic" curvature:

> We discussed how curvature does not happen within the dimensions the object is in, but in a new such that we cannot observe that space itself gets curved. (Student group 8)

This response reflects a common criticism brought forward by physicists and physics educators (e.g. Gould, 2016), namely that the RSA suggests that spacetime curves into an unseen additional dimension. Indeed, it seems that students struggled with this depiction of spacetime and were not necessarily aware that the unseen dimension is an artefact of the analogy that does not correspond to a real physical phenomenon.

The second most common analogical weakness identified by students was the problem related to the force of gravity:

> Large masses curve everything of time and space, and do this in several dimensions. This can be difficult to visualize, so we usually look at this in two dimensions. Then it looks as if there is something pulling the mass. But it is the mass itself that curves space. (Student group 9)

Here, students summarized the key problem of the RSA addressed in the cartoon: that "something" is needed to exert a pull on the massive object. In the source domain of the rubber sheet, this pull is provided by the force of gravity – which does not have an analogue in the target domain of abstract spacetime.

Almost all of the ten responses tagged by the code "force of gravity" addressed the incorrect assumption that the mass is being pulled down:

> We discussed that this model is a bit wrong to use, because it refers to a force of gravity that holds the ball down. It does not work like this according to Einstein. . (Student group 10)

However, student discussions about the analogy suggest that many still thought along the lines of classical physics or struggled with reconciling how masses can exert an influence without the mediating force of gravity:

---

[1] The original Norwegian „å se for seg“ can be translated as „to visualise“, „to envision“, „to see in your mind's eye“ or more literally "to see in front of you". In our translations we chose the expression „to visualize“.

> *It is difficult to describe spacetime. It is also difficult to visualize that mass influences spacetime just by being mass. That mass is not influenced by a force and that that's the reason it exerts an influence.* (Student group 11)

> *We discussed how mass can influence spacetime, if spacetime does not have mass itself. (*Student group 12*)*

This finding shows us that, for students, the phenomenon of gravity seems to be deeply associated with the concepts of force and mass. In particular, the second quote is interesting, as it expresses the idea that spacetime itself might have mass in order for it to be influenced by other masses. This finding gives insight into students' ability to use their existing knowledge to deduce characteristics of novel scientific concepts. In this case, the justified conclusion that spacetime must have mass because it reacts to the presence of other masses was discarded by the students themselves.

Another important feature that the RSA hides is that spacetime has a temporal dimension - masses curve time as well. Even though the learning environment introduced the RSA by pointing out that this is a weakness of the analogy, only few students addressed this weakness:

> *The first analogy does not take the time coordinate into account and there is a simplified model. Therefore, there arise questions concerning imprecisions of the analogy. We have to use simplified models because we cannot visualize four dimensions.* (Student group 13)

> Student group? *We discussed Einstein's model where curvature in spacetime and geometry around it lead to what we call gravity. The most difficult to understand is the time parameter in the model and how also this is curved.* (Student group 14)

The relatively few responses that mentioned the time dimension suggest that, generally, students were not aware that the time dimension plays an important role in the origin of gravitation; those that showed such awareness admitted that this part of the theory was difficult to understand. This observation suggests that the role of time might have posed a conceptual challenge for students when dealing with gravity in the setting of GR. Alternatively, the cartoon might have set students on a different track by emphasizing the force aspect of the analogy, making them neglect the time aspect.

Interestingly, students usually only addressed one flaw of the RSA. Of the 26 responses that addressed a specific strength or weakness, only four mentioned two specific flaws/strengths and three of those responses mentioned the dimension problem. Even though the task was open-ended, thus allowing students to explore different problems that come with the use of the RSA, most seemed to have settled on one problematic issue. This observation suggests that there might be instructional potential to facilitate conceptual understanding of GR by presenting various strengths and weaknesses of the RSA explicitly.

The last theme, "cartoon context", does not directly relate to the RSA, but offers interesting insights into students' conceptualization of gravity in GR nonetheless. The theme contains the codes "teaching situation" and "Newton/Einstein".

Many students picked up on the teaching situation illustrated in the cartoon that emphasized the role of the teacher when learning GR. They stated that it is difficult to teach GR and that falling back to mathematics might be a convenient way for teachers to avoid facing difficult questions by students:

> *We discussed that the teacher didn't have a good response to the question of the student and responded with a really theoretical calculation to stop the questions. The reason for this can be that it is impossible for us to visualize four dimensions, and therefore it is also difficult to teach this.* (Student group 15)

> *First the teacher explains via drawings, the student does not understand this, so it gets explained via formulas and logic and the student thinks this is boring. The topic is possibly also too difficult for the student to understand if you just jump right into it.* (Student group 16)

Surprisingly, interpretations of the teaching situation produced an interesting response in five cases: students compared the teacher to Einstein and the student to Newton. This is in line with the presentation of GR in the program where GR is presented in opposition to Newtonian physics. Students projected the Newtonian and the Einsteinian view on the two protagonists in the cartoon. Here, we recognize an observation made already during the metaphor analysis: students drew on previous presentations of GR in the program and several students seemed to remember the contrast between Einstein and Newton well:

> *The teacher is Einstein, while the student is Newton.* (Student group 17)

> *First the teacher tries to explain how time and space can be curved by objects. The student doesn't understand this and he tries to explain it with equations instead. He thinks this is boring. These are two persons that maybe have two different ways to look at spacetime. The teacher looks at it in the same way as Einstein and the student in the same way as Newton. Therefore, they don't quite understand each other.* (Student group 18)

We have used the thematic analysis of student responses to explore student awareness of the analogical nature of the RSA and to answer our third research question. In summary, we can see that students displayed a sound understanding of the scope and limitations of analogies as one particular model in the domain of GR. Nonetheless, students addressed specific strengths and weaknesses less often. The reduction of the number of dimensions and the incorrect mechanism of the curving of the rubber sheet by means of the classical force of gravity were the weaknesses that students mentioned most. Less common was the observation that the RSA only depicts curved space and thus neglects the curvature of the time component in spacetime.

5. Discussion

We began with the goal of understanding the RSA and the affordances it provides for students to conceptualize gravity as curved spacetime in the domain of GR. In this section, we want to summarize our findings in light of our research questions and discuss instructional implications related to the approach of embodied cognition.

Two rounds of independent analyses of student responses (coding for conceptual metaphors and coding for strengths and weaknesses of the RSA) showed that students generated more conceptual metaphors than the ones found in the literature. The greater part of the conceptual metaphors had much overlap with the ones employed by experts in the field and merely deconstructed spacetime into its space and time components. However, we observed novel mappings between the target and source domains as well, and those mappings led to essentially different conceptions: whereas GR posits that force is a consequence of the curvature of spacetime (we interpret geometrical properties as forces acting on objects), students turned this reasoning upside down. They described a force that curves spacetime or talked about gravity curving spacetime. Thus, metaphor theory suggests that students might confuse cause and effect when working with the RSA.

Niebert and Gropengießer (2014) made a related observation concerning students' conceptions of the greenhouse effect. They found that students and scientists used the same source and target schemata but mapped them differently, leading to different conceptions of the greenhouse effect. Selecting those mappings that will be fruitful when conceptualizing scientific concepts is thus an intricate task in abstract domains such as climate change or general relativity.

We casted our investigations into the framework of embodied cognition, which assumes that conceptual understanding requires grounding in experience (Niebert et al., 2012). According to this framework, it is not enough to relate instructional analogies to everyday life. Students use their embodied experience to understand analogies, something that instructors need to be aware of. For analogies to be successful in communicating scientific concepts, the chosen source domains need to be embodied in such a way as to not conflict with the target domain. A metaphor for gravity should not depend on student's embodied experiences with gravity.

In light of our findings, we would like to put this observation further into perspective. Even though the source domain of the RSA draws on students' embodied experience, it seems that exactly this conceptualized experience of gravity often got in the way of inferring the right analogical mappings. In order to conceptualize the physical mechanism of gravity in the domain of GR, students need to develop awareness of the tension between the physical force of gravity in the everyday experiential sense and the curved spacetime explanation.

Nonetheless, the RSA gives students a concrete object to visualize and interact with. If we make students become aware of the scope and limitations of their imaginative capacities, this analogical visualization could fill in a link in the chain of reasoning leading from experiential understanding of gravity towards a more sophisticated understanding in the context of GR. After all, many students linked their understanding of curved spacetime to their ability to visualize it. This finding resonates with a shared interest in visualizations among science educator who have called attention to the significance of developing students' skills of visualization more systematically (Gilbert, 2005).

More generally, we argue that GR is a domain in which students can benefit from a teaching approach with a greater emphasis on the nature of science and scientific models, in particular on the scope and limitations of scientific models. While many students displayed a good understanding of the role that analogies and models play in GR, significantly fewer identified specific limitations of the RSA. There seems to be untapped potential in creating awareness for exactly those misleading features of the RSA in order to foster conceptual understanding of relativistic phenomena. We have thus identified several specific instructional strategies for improving the introduction of GR in classroom settings. First, we suggest that teachers might provide an explicit classroom discussion of the flaws of the RSA as listed in Table 2. Identifying the shortcomings of a two-dimensional, spatial representation of four-dimensional curved spacetime can help prevent the formation of mismatches and incorrect mappings between target and source domains.

The RSA is one way of visualizing the physics of curved spacetime. To prevent the one-sided presentation of the concept of curved spacetime as a deformed rubber sheet, teachers can supplement this analogy with other models of spacetime such as the world map model that compares the geometry of spacetime to the geometry of two-dimensional maps (Gould, 2016; Stannard et al., 2017). Seeing that the time dimension tends to be a neglected feature in the RSA, it is moreover important to emphasize that curvature and movement in spacetime entails both curvature and movement in space and time. The role of time as a crucial part of teaching and instructional in GR is taken up in a related study of project ReleQuant (Steier & Kersting, n.d.). In this case study, that reports on the first trial of the ReleQuant project learning environment, students struggled to use Einstein's model, and in particular the RSA, to explain gravitational phenomena from everyday life. While they could explain planetary movement according to GR, students failed to draw on Einstein's model to explain why they were pulled towards the ground. It seemed that students related curvature to movement and lacked an understanding of their continuous movement along the time-dimension. Teachers should thus pay particular attention to the role of time when using the RSA to teach GR.

In their discussions, students frequently juxtaposed Newton's explanations of gravity to Einstein's and identified the stickmen in the cartoon with Newton and Einstein respectively. Thus, another fruitful way for teachers to introduce the physics of GR might be to address the historic development of GR and Einstein's struggle to overcome Newtonian physics. Helping students contrast their own classical conceptions of gravity with the novel relativistic ones can serve as a fruitful addition to the use of the RSA.

More generally, linking the concept of spacetime and other key concepts of GR to students' life worlds is one design principle for learning resources that project ReleQuant has identified as important in the domain of GR (Kersting et al., 2018). To counteract the lack of experience with relativistic phenomena, visualizations in form of digital simulations and animations can supplement static representations of spacetime.

In this study, students encountered the analogy as part of a learning sequence that guided them through different explorations of curved spacetime in form of interactive simulations. Each separate task provided a slightly different perspective on spacetime, which constitutes one conceptually important part of GR. The way Einstein modeled gravitational phenomena through geometric reasoning extended his original ideas about the principle of relativity and the relation of space and time in special relativity. A broader account of this development and students' understanding of other concepts in GR is given in (Kersting et al., 2018).

Finally, we would also like to discuss what we view as two important limitations of this study. First, our data consist of written responses from five physics classes in three Norwegian upper secondary schools. The analysis of students' metaphorical language allowed us to gain insight into the ways that students conceptualised curved spacetime through the RSA. These insights are, however, often only supported by a small number of responses. While our results are thus not generalizable per se, we think that knowledge of the student-generated mappings between target and source domain can help teachers to identify possible sources of conflict with the RSA. This knowledge has thus the potential to be quite broadly applicable in teaching and instruction of GR.

Second, the discussion task featuring the cartoon of the RSA was open-ended. Students were thus not necessarily interpreting the cartoon in a way that aligned with our research questions. Asking students to use their knowledge of GR to discuss and comment on the cartoon can of course only give a partial insight into their conceptual understanding of gravity and curved spacetime. Keeping this in mind and viewing our study as a first step towards a more holistic understanding of learning processes in GR, however, the format of the task had advantages as well: the cartoon addressed explicitly a particular flaw of the RSA, thus prompting students to comment on its analogical nature. Also, by leaving the task open, students could not merely repeat back answers from other parts of the learning environment – a common behavior that we had observed in the first trialling of the learning environment.

6. Conclusion

Addressing the controversy around the use of the RSA in the teaching and learning of GR, this study presents empirical evidence of upper secondary school students' reasoning when conceptualizing gravity as curved spacetime. First, we performed a metaphor analysis of the literature to identify four conceptual metaphors that comprise the fundamental relationships between target and source domains of the RSA. Based on this analysis, we identified strengths and weaknesses of the RSA. A second metaphor analysis of students' written responses revealed a greater variety of student-generated conceptual metaphors than held by experts in the field and a thematic analysis gave insight into students' awareness of the analogical nature of the RSA.

We hope that knowledge of students' different conceptions of gravity as curved spacetime and our compilation of strengths and weaknesses of the RSA can give guidance for teachers and science educators alike. Making students become aware of the strengths and weaknesses of the RSA might be as important as introducing them to the physics of gravitation according to the relativistic framework. Moreover, teaching should be explicit about identifying the source and target domains in order to make it clearer to students what a metaphor is and how it is used. One area to be explored in future work would be to teach students simultaneously about gravity with the RSA along with an introduction to the structural features of metaphors more generally so that they can better interpret and apply the RSA.

Moreover, our study contributes to a growing body of recent research on metaphor analysis and embodied cognition in the field of science education. While previous research has found that it takes more than connecting analogies and metaphors to students' everyday life, namely an analogy that employs embodied sources (Niebert et al., 2012), we present findings that give important nuances to this observation. We observed a conflict between students' embodied understanding of gravity and the abstract description of GR. Even though the source domain of the RSA draws on students' embodied experience, it seems that exactly this conceptualized experience of gravity can get in the way of inferring the right analogical mappings. Thus, even if an analogy builds on an embodied source domain, it can fail in communicating scientific concepts fruitfully. It is therefore crucial that students probe their imaginative skills when conceptualizing abstract scientific concepts such as curved spacetime to build awareness for the processes of their own metaphorical reasoning.

Despite some inherent conceptual flaws, the RSA has the potential to serve as a good metaphor. Teaching GR can be successful if approaches build on students' understanding of the limited nature of scientific models, communicate explicitly target and source domain and strengths and weaknesses of the RSA, and point out the disagreement between students' experiential understanding of gravity and the reliance of the RSA on exactly this experiential understanding to explain gravity in more abstract terms.

Compliance with Ethical Standards

Conflict of Interest: No potential conflict of interest was reported by the authors.

Acknowledgements

This work was supported by the Research Council of Norway (ProjectNo. 246723) and the Olav Thon Foundation.

References

Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M. R., Acernese, F., Ackley, K., … Zweizig, J. (2016). Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters, 116*(6), 061102.

Amin, T. G., Jeppsson, F., & Haglund, J. (2015). Conceptual Metaphor and Embodied Cognition in Science Learning: Introduction to special issue. *International Journal of Science Education*, *37*(5–6), 745–758.

Aubusson, P. J., Harrison, A. G., & Ritchie, S. M. (2006). *Metaphor and Analogy in Science Education* (Vol. 30). Springer.

Baldy, E. (2007). A New Educational Perspective for Teaching Gravity. *International Journal of Science Education*, *29*(14), 1767–1788.

Bandyopadhyay, A., & Kumar, A. (2010a). Probing students' ideas of the principle of equivalence. *European Journal of Physics*, *32*(1), 139–159.

Bandyopadhyay, A., & Kumar, A. (2010b). Probing students' understanding of some conceptual themes in general relativity. *Physical Review Special Topics - Physics Education Research*, *6*(2), 020104.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3,* 77–101.

Callin, N. P., Pålsgård, J., Stadsnes, R., & Tellefsen, C. W. (2012). *ERGO Fysikk 2*. Oslo: Aschehoug.

Casey, E. S. (1979). *Imagining*. Bloomington: Indiana University Press.

Chandler, M. (1994). Philosophy of Gravity : Intuitions of Four- Dimensional Curved Spacetime. *Science & Education*, 155–176.

Chen, Y.-C., Park, S., & Hand, B. (2016). Examining the Use of Talk and Writing for Students' Development of Scientific Conceptual Knowledge Through Constructing and Critiquing Arguments. *Cognition and Instruction*, *34*(2).

Dimitriadi, K., & Halkia, K. (2012). Secondary Students' Understanding of Basic Ideas of Special Relativity. *International Journal of Science Education, 34*(16), 2565–2582.

diSessa, A. A. (1981). An elementary formalism for general relativity. *American Journal of Physics*, *49*(1981), 401.

Einstein, A. (1915). Grundgedanken der allgemeinen Relativitätstheorie und Anwendung dieser Theorie in der Astronomie [Fundamental Ideas of the General Theory of Relativity and the Application of this Theory in Astronomy]. *Preussische Akademie Der Wissenschaften, Satzungsberichte*, 315.

Einstein, A. (1917). *Über die spezielle und die allgemeine Relativitätstheorie [Relativity: The Special and General Theory]*. Braunschweig: Vieweg.

Farr, B., Schelbert, G., & Trouille, L. (2012). Gravitational wave science in the high school classroom. *American Journal of Physics, 80*(10), 898.

Gentner, D., Bowdle, B., Wolff, P., & Boronat, C. (2001). Metaphor is like analogy. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 199–253). Cambridge, MA: MIT.

Gilbert, J. (2004). Models and Modelling: Routes To More Authentic Science Education. *International Journal of Science and Mathematics Education*, *2*(2), 115–130.

Gilbert, J. (Ed.). (2005). *Visualization In Science Education*. *Visualization in Science Education Vol. 1*. Dordrecht: Springer.

Gould, R. R. (2016). Why does a ball fall?: A new visualization for Einstein's model of gravity. *American Journal of Physics*, *84*(5), 396–402.

Greene, B. (2010). *The Elegant Universe*. New York: W W Norton & Co Inc.

Haglund, J. (2017). Good Use of a 'Bad' Metaphor - Entropy as Disorder. *Science & Education*, 205–214.

Harrison, A. G., & Treagust, D. F. (2006). Teaching and learning with analogies : friend or foe. In A. G. Harrison & S. M. Ritchie (Eds.), *Metaphor & Analogy in Science Education* (pp. 11–24). Netherlands: Springer.

Hartle, J. B. (2005). General Relativity in the Undergraduate Physics Curriculum. *American Journal of Physics*, *14*(2006), 9.

Henriksen, E. K., Bungum, B., Angell, C., Tellefsen, C. W., Frågåt, T., & Vetleseter Bøe, M. (2014). Relativity, quantum physics and philosophy in the upper secondary curriculum: challenges, opportunities and proposed approaches. *Physics Education*, *49*(6), 678–684.

Hentschel, K. (Ed.). (1998). *The Collected Papers of Albert Einstein, Volume 8 (English) The Berlin Years: Correspondence, 1914-1918*. Princeton, New Jersey: Princeton University Press.

Hesse, M. (1952). Operational Definition and Analogy in Physical Theories. *British Journal for the Philosophy of Science*, 281–94.

Hesse, M. (1953). Models in Physics. *British Journal for the Philosophy of Science*, *4*, 98–214.

Jerstad, P., Sletbak, B., Grimenes, A. A., Renstrøm, R., Holm, O. B., & Nymo, M. (2014). *RomStoffTid Fysikk 2*. Oslo: Cappelen Damm.

Kampourakis, K. (2016). The Bad Use of Metaphors and the Use of Bad Metaphors. *Science & Education*, *25*(9–10), 947–949.

Kapon, S., & DiSessa, A. A. (2012). Reasoning Through Instructional Analogies. *Cognition and Instruction*, *30*(3), 261–310. http://doi.org/10.1080/07370008.2012.689385

Kaur, T., Blair, D., Moschilla, J., Stannard, W., & Zadnik, M. (2017a). *Teaching Einsteinian Physics at Schools: Models and Analogies - Part 1 (Manuscript draft)*. Retrieved from http://arxiv.org/abs/1704.02058

Kaur, T., Blair, D., Moschilla, J., Stannard, W., & Zadnik, M. (2017b). Teaching Einsteinian physics at schools: Part 3, review of research outcomes. *Physics Education*, *52*(6). http://doi.org/10.1088/1361-6552/aa83dd

Kersting, M., Henriksen, E. K., Bøe, M. V., & Angell, C. (2018). General relativity in upper secondary school: design and evaluation of an online learning environment using the model of educational reconstruction. *Physical Review Physics Education Research*, *14*(1), 010130-1-010130-18. http://doi.org/10.1103/PhysRevPhysEducRes.14.010130

Kind, P. M., & Kind, V. (2007). Creativity in Science Education: Perspectives and Challenges for Developing School Science. *Studies in Science Education*, *43*(1), 1–37. http://doi.org/10.1080/03057260708560225

Krijtenburg-Lewerissa, K., Pol, H. J., Brinkman, A., & Van Joolingen, W. R. (2017). Insights into teaching quantum mechanics in secondary and lower undergraduate education: A literature review. *Physical Review Physics Education Research*, *13*.

Lakoff, G., & Johnson, M. (2003). *Metaphors We Live By* (2nd ed.). Chicago: University of Chicago Press.

Lancor, R. (2014a). Using Metaphor Theory to Examine Conceptions of Energy in Biology, Chemistry, and Physics. *Science & Education*, *23*(6), 1245–1267.

Lancor, R. (2014b). Using Student-Generated Analogies to Investigate Conceptions of Energy: A multidisciplinary study. *International Journal of Science Education*, *36*(1), 1–23.

Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Norwood, New Jersey: Ablex Publishing Corporation.

Levrini, O. (2014). The role of history and philosophy in research on teaching and learning of relativity. In M. R. Matthews (Ed.), *International Handbook of Research in History, Philosophy and Science Teaching* (pp. 157–181). Dordrecht: Springer.

Levrini, O., & DiSessa, A. A. (2008). How students learn from multiple contexts and definitions: Proper time as a coordination class. *Physical Review Special Topics - Physics Education Research*, *4*(1), 1–18. http://doi.org/10.1103/PhysRevSTPER.4.010107

Mach, E. (1893). *The Science of Mechanics*. (Thomas Joseph McCormack, Ed.). Open Court Publishing Company.

Middleton, C. A., & Weller, D. (2016). Elliptical-like orbits on a warped spandex fabric: A theoretical/experimental undergraduate research project. *American Journal of Physics*, *84*(4), 284–292.

Nemirovsky, R., Kelton, M. L., & Rhodehamel, B. (2012). Gesture and imagination On the constitution and uses of phantasms. *Gesture*, *12*(2), 130–165.

Niebert, K., & Gropengießer, H. (2014). Understanding the Greenhouse Effect by Embodiment – Analysing and Using Students' and Scientists' Conceptual Resources. *International Journal of Science Education*, *36*(2), 277–303.

Niebert, K., Marsch, S., & Treagust, D. F. (2012). Understanding Needs Embodiment : A Theory-Guided Reanalysis of the Role of Metaphors and Analogies in Understanding Science. *Science Education*, 849–877.

Pitts, M., Venville, G., Blair, D., & Zadnik, M. (2014). An Exploratory Study to Investigate the Impact of an Enrichment Program on Aspects of Einsteinian Physics on Year 6 Students. *Research in Science Education*, *44*(3), 363–388.

Poincaré, H. (1898). La mesure du temps [The measure of time]. *Rev. Metaphys. Morale*, *6*(1).

Price, R. H. (2016). Spatial curvature, spacetime curvature, and gravity. *American Journal of Physics*, *84*(8), 588–592.

Rasmussen, I., & Ludvigsen, S. (2010). Learning with computer tools and environments: A sociocultural perspective. In C. W. & J. K. S. K. Littleton (Ed.), *International Handbook of Psychology in Education* (pp. 399–435). UK, Emerald.

Reichenbach, H. (1928). *Philosophie der Raum-Zeit-Lehre [The Philosophy of Space and Time]*. New York: Dover Publications.

Roth, W. M., & Lawless, D. (2002). Science, Culture, and the Emergence of Language. *Science Education*, *86*(3), 368–385.

Russell, B. (1925). *ABC of Relativity*. (Kegan-Paul, Trench, & Trubner, Eds.). London: The ABC of Relativity.

Sartre, J.-P. (2004). *The imaginary. A phenomenological psychology of the imagination.* New York: Routledge.

Schmitt, R. (2005). Systematic Metaphor Analysis as a Method of Qualitative Research. *The Qualitative Report*, *10*(2), 358–394.

Schön, D. A. (1979). Generative metaphor: A perspective problem-setting in social policy. In A. Ortony (Ed.), *Metaphor and thought* (pp. 254–283). Cambridge, England: Cambridge University Press.

Scott, P., & Mortimer, E. (2005). Meaning making in high school science classrooms: a framework for analysing meaning making interactions. *Research and the Quality of Science Education*, *7*, 395–406.

Silva, C. C. (2007). The Role of Models and Analogies in the Electromagnetic Theory: A Historical Case Study. *Science & Education*, *16,* 835. http://doi.org/https://doi.org/10.1007/s11191-006-9008-z

Stadermann, H. K. E., & Goedhart, M. J. (2017). Comparison and analysis of quantum physics in secondary school curricula of 13 different countries. Presentation at GIREP 2017, Dublin, Ireland.

Stannard, W., Blair, D., Zadnik, M., & Kaur, T. (2017). Why did the apple fall? A new model to explain Einstein’s gravity. *European Journal of Physics*, *38*(1), 015603.

Steier, R., & Kersting, M. (n.d.). Metaimagining and embodied conceptions of spacetime. *Manuscript under Review*.

Stinner, A. (2003). Scientific Method, Imagination and the Teaching of Physics. *Physics World*, 335–346.

The Norwegian Directorate for Education and Training. (2006). Physics - programme subject in programmes for specialization in general studies. Retrieved from http://www.udir.no/kl06/fys1-01/Hele/Kompetansemaal/fysikk-2/?lplang=eng

Thorne, K. (2009). Warping Spacetime. In G. W. Gibbons, E. P. S. Shellard, & S. J. Rankin (Eds.), *The Future of Theoretical Physics and Cosmology: Celebrating Stephen Hawking’s Contributions to Physics*. Cambridge: Cambridge University Press.

Treagust, D. F., & Duit, R. (2015). On the Significance of Conceptual Metaphors in Teaching and Learning Science: Commentary on Lancor; Niebert and Gropengiesser; and Fuchs. *International Journal of Science Education*, *37*(5–6), 958–965.

Vailati, E. (1997). *Leibniz and Clarke: A Study of Their Correspondence*. Oxford and New York: Oxford University Press.

Velentzas, A., & Halkia, K. (2013). The Use of Thought Experiments in Teaching Physics to Upper Secondary-Level Students: Two examples from the theory of relativity. *International Journal of Science Education*, *35*(18), 3026–3049.

Viereck, G. S. (1929). What Life Means to Einstein. *The Saturday Evening Post*. http://doi.org/http://www.saturdayeveningpost.com/wp-content/uploads/satevepost/what_life_means_to_einstein.pdf

Vygotsky, L. (1962). *Thought and language.* Cambridge: M.I.T. Press.

Watkins, T. R. (2014). *Gravity & Einstein: Assessing the Rubber Sheet Analogy in Undergraduate Conceptual Physics (Master Thesis)*. Boise State University.

Zahn, C., & Kraus, U. (2014). Sector Models-A Toolkit for Teaching General Relativity. Part 1: Curved Spaces and Spacetimes. *European Journal of Physics*, *35*, 055020.