

## Developing conceptual understanding in ray optics via learning with multiple representations

Rosa Hettmannsperger · Andreas Mueller · Jochen Scheid · Wolfgang Schnotz

Published online: 30 September 2015  
© Springer Fachmedien Wiesbaden 2015

**Abstract** A particular difficulty in physics learning is the fact that pupils' "intuitive" concepts are often resistant to instruction. This article reports empirical results from two related studies within an interdisciplinary project of physics education and educational psychology in ray optics. Two different kinds of treatment groups (TG A and TG B), both targeted at widespread pupils' intuitive concepts ( $N=511$ ), were compared with the results of a control group (CG C) learning with conventional tasks ( $N=218$ ) provided by a related study II. Pupils in TG A) of study I worked on cognitively activating tasks related to widespread intuitive concepts in ray optics explicitly requiring them to deal with multiple representations. Pupils in the TG B) of study I worked on the same intuitive concepts, but without the cognitively acti-

---

Dr. R. Hettmannsperger (✉)  
Institut Universitaire de Formation des Enseignants (IUFE), Université de Genève,  
Boulevard du Pont d'Arve 40,  
1211 Genève, Switzerland  
e-mail: rosa.hettmannsperger@unige.ch

Prof. Dr. A. Mueller  
Faculté des Sciences/Section Physique,  
Université de Genève/Boulevard du Pont d'Arve 40,  
1211 Genève, Switzerland  
e-mail: andreas.mueller@unige.ch

Dr. J. Scheid  
Institut für naturwissenschaftliche Bildung (InB),  
Universität Koblenz-Landau, Fortstraße 7,  
76829 Landau, Germany  
e-mail: scheid@uni-landau.de

Prof. Dr. W. Schnotz  
Faculty of Psychology, University of Koblenz-Landau,  
Fortstraße 7,  
76829 Landau, Germany  
e-mail: schnotz@landau.de

vating representational component. TG A) and B) were compared with each other and with CG C) learning with conventional tasks. The results indicated that tasks addressing widespread intuitive pupils' concepts improved conceptual understanding significantly more than conventional tasks. There was evidence of a significant intermediate effect showing medium-term stability.

**Keywords** Conceptual understanding · Learning with multiple representations · Ray optics

## **Förderung von konzeptuellem Verständnis in der Strahlenoptik via Lernen mit multiplen Repräsentationen**

**Zusammenfassung** In diesem Artikel werden Forschungsergebnisse aus zwei aufeinander bezogenen Studien unter dem Gesichtspunkt des konzeptuellen Verständnisses in der Strahlenoptik verglichen. Zwei verschiedene Treatmentgruppen einer Studie I, in denen jeweils Schülervorstellungen thematisiert wurden ( $N=511$ ), wurden mit einer Kontrollgruppe (KG C) einer zweiten Studie verglichen, in der die Schüler/innen konventionelle Aufgaben bearbeiteten ( $N=218$ ). Schüler/innen der TG A) aus Studie I bearbeiteten kognitiv aktivierende Aufgaben, welche verbreitete Schülervorstellungen in der Strahlenoptik thematisierten, wobei sie kognitiv aktiviert wurden sich explizit mit multiplen Repräsentationen zu befassen. Schüler/innen der TG B) aus Studie I setzten sich exakt mit den gleichen Schülervorstellungen auseinander jedoch ohne kognitive Aktivierung in Bezug auf das Lernen mit multiplen Repräsentationen. Beide Gruppen aus Studie I wurden miteinander sowie mit der KG C) aus Studie II verglichen. Die Ergebnisse belegen, dass das Thematisieren von weitverbreiteten Schülervorstellungen im Vergleich zu konventionellen Aufgaben das konzeptuelle Verständnis mittelfristig, signifikant mit mittlerer Effektstärke verbessert.

**Schlüsselwörter** Konzeptuelles Verständnis · Lernen mit multiplen Repräsentationen · Strahlenoptik

### **1 Introduction and research questions**

A growing body of research deals with the representational aspects of developing learners' conceptual learning and change, pointing out that students need to develop and understand multiple representations to improve their understanding of basic scientific concepts (Plötzner and Spada 1998; Wilhelm 2005; Botzer and Reiner 2005; Sell et al. 2006; Waldrip et al. 2006; Mortimer and Buty 2009; Taber 2009; Hubber et al. 2010; Tsui and Treagust 2013).

Here, we understand multiple representations (MRs) as various perceptually and/or conceptually different forms of external and internal mental representations, which serve as cognitive tools and support understanding as well as the use of abstract concepts (Ainsworth 2008; Gilbert and Treagust 2009; Tsui and Treagust 2013).

An essential point in science learning is that a proper understanding of the relationship between scientific experiments and phenomena, on the one hand, and of their conceptual basis, on the other hand, requires the learner to use MRs at different levels of abstraction. Examples of their use are describing experimental setups, procedures and observations with oral or written language in terms of the appropriate concepts, using related schematic or logical images such as ray or vector diagrams, or expressing experimental results with functional graphs or mathematical relations containing formal representations of the concepts in question.

Throughout various fields of science education, the importance of MRs for conceptual learning has been well established, notably in chemistry (Taber 2009), in particular for the fundamental topic of micro-macro level connection (Cheng and Gilbert 2009), in biology (Tsui and Treagust 2013) and in geo-sciences (Sell et al. 2006). For instance, in physics, Hubber et al. (2010) proved the potential of conceptual learning through multiple representations for the area of mechanics.

With regard to conceptual learning from and about experiments, cognitive conflicts with discrepant experience is considered a standard approach for fostering conceptual understanding, both for a scientific discipline (Thagard 1991) and for individual learners (Kim et al. 2002; Lee et al. 2003; Zimrot and Ashkenazi 1991). While the theoretical foundation of conceptual change is still under discussion (Galili and Hazan 2000; Schnotz 2006; Özdemir and Clark 2007), there is a strong consensus that a crucial element for conceptual change is the fact that learners deal actively with discrepant information. In other words, learners need to use their own cognitive and representational resources, as emphasized also by Hubber et al. (2010).

One strategy for engaging pupils to learn with MRs is the use of cognitively activating tasks (Lipowsky 2009, p. 93). In the present context, “cognitive activation” means that learners think more often, more explicitly, and more deeply about experiment-related representations, express and draw conclusions from them. One might say this is the case in a typical learning setting as well. However, the type of learning activity for achieving cognitive activation is new in the sense that specific tasks have been developed that require explicit reasoning about various experiment-related representations.

In the context of representation related cognitive activation aiming at conceptual change, we investigated the following questions:

1. Do tasks aiming at fostering conceptual change based on cognitive activation through MR learning (referred to as A) help pupils to develop a deeper conceptual understanding compared with tasks dealing with the same variety and number of intuitive concepts, but without specific MR-based cognitive activation (referred to as B)?
2. In order to have a baseline value, we asked additionally: Are both of these kinds of tasks A) and B) more effective than learning with conventional tasks C)?

## 2 Theoretical and empirical background

In order to understand how representation related cognitive activation might help learners to develop conceptual understanding, we have to analyze how learners' cognitive processes are related to the information presented at school.

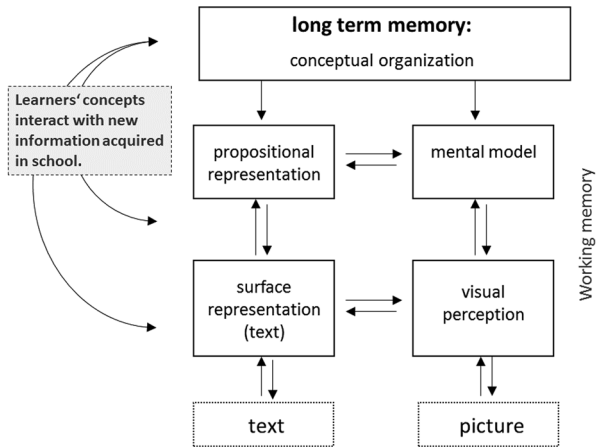
From a cognitive point of view, learners in physics classrooms deal mainly with two kinds of information sources: first, demonstrations or students' experiments that allow them to observe or experience phenomena in real life and second, representations. The latter are objects or events that stand for the phenomenon in physics, or for a related model explaining the observed phenomenon.

With regard to processes of teaching and learning, we can distinguish between representations from at least three perspectives (see Ainsworth 2006). First, representations can differ in terms of their format; there are descriptive representations consisting of symbols and there are depictive representations consisting of icons (Schnotz 2005, p. 52). Depictive representations either represent an object or a process in a realistic manner including correspondence in terms of physical properties or in a more abstract manner displaying a structure that is analog to the subject matter. An example for a realistic representation could be a photo of an experimental set-up using a convex lens to obtain a real image of a shining object and for an abstract representation the related ray diagram that represents the construction of the image and its size. The latter representation can be also characterized as depictive schematic logic representation. Descriptive representations can be verbal or mathematical-numeric or mathematical-symbolic (e.g., the equation for calculating the scale of magnification or reduction of a real image in ray optics). Second, representations differ in their role within teaching and learning: For example, they can be presented externally on a paper or a screen or exist internally in teachers' or learners' minds (Cox 1999, p. 344; Schnotz and Bannert 2003, p. 143). Third, different kinds of representation are often ranked according to their level of abstraction (Leisen 1998, p. 9) e.g., the photo of the experimental setup is less abstract than a ray diagram.

According to the theoretical framework for analyzing text and picture comprehension (Schnotz and Bannert 2003; Schnotz 2005), when a learner reads a text about new learning content, for example in physics, listens to the teachers' explanations or looks at an image, first of all, he or she forms a mental representation of the surface structure, see Fig. 1. After that, a verbal or a visual pictorial filter selects the information and forwards it to working memory. Here the verbal information leads to a propositional representation and later to the construction of a mental model. Visual pictorial information, on the other hand, leads directly to the construction of a mental model. They involve task-based selection of information obtained from the long-term memory in order to process information in the working memory (Dutke 1994, p. 76 f.) and help learners to understand an issue by simulating processes or envisioning information internally. During learning processes, mental models and propositional representations interact continuously with each other. On the one hand, the construction of mental models can be based on propositional representations guided by cognitive concepts from long-term memory. On the other hand, new propositional information can be read off from mental models in working memory and added to the former representations in long-term memory.

Learners' concepts in long-term memory exert an influence on the construction of mental models and surface representations as well as on visual perception (Schnotz and Bannert 2003, p. 145; Schnotz 2006, p. 75 f.); as shown in Fig. 1. Within this process, intuitive concepts play a decisive role, as they can lead to incorrect conclu-

**Fig. 1** Integrated model of text and picture comprehension (according to Schnotz and Bannert 2003, p. 145; modified by Hettmannsperger, Müller, Scheid and Schnotz)



sions when solving a task (due to inadequate mental models) and they can lead to (new) problematic concepts in the long-term memory derived from mental models.

In addition, learners' internal representations and mental models influence the external representations created while solving tasks (Gentner and Gentner 1983, p. 118; Cox 1999, p. 354). Especially in physics, problem-solving regularly requires learners to identify or develop at least one appropriate representation of a given problem or task. To solve the problem, they then continue to operate on dealing with the representation(s) they have identified or developed. If they need several representations, this process requires learners to be systematically engaged in mutually relating different kinds of MRs at different levels of abstraction or in different formats (Ainsworth 1999, p. 142, 2006, p. 6).

While taking into consideration the mutual relations between widespread intuitive pupils' concepts, external representations, and the construction of mental models, we can refer to the empirical results of two recent strands of research: one is learning with MRs from a cognitive psychological point of view, as outlined above, and the other is fostering conceptual understanding while dealing with MRs in science classrooms, to be outlined in the following section.

In recent years, various researchers in physics education have pointed out that dealing with MRs helps learners to improve their understanding of basic scientific concepts.

Plötzner and Spada (1998, p. 81 f.) analyzed how qualitative representations of problems can be used to solve quantitative problems in the field of Newton's mechanics at secondary level II. Their results indicated that the number of correctly solved tasks decreased significantly when students believed in the impetus<sup>1</sup> concept.

Cheng and Shipstone (2003, p. 193 f.) investigated a new approach for teaching the electric circuit theory at secondary level II using different kinds of diagrams illustrating how current, voltage, resistance, and power are distributed. The authors explained how these representations supported students' learning of basic electrical concepts and provide them with useful strategies for solving both quantitative and qualitative problems.

Wilhelm (2005, p. 175–216) compared a treatment group that dealt with representations of two-dimensional movements on different levels of abstraction with a control group that intensively learned to interpret one-dimensional movements represented in graphs. Students in the treatment group outperformed students in the control group in tasks about the direction of acceleration and two-dimensional movements and understood concepts about directions more often than students in the control group.

Botzer and Reiner (2005) focused on analyzing the representational requirements for acquiring scientific concepts in the field of magnetism at secondary level I about magnetic fields. Students completed a number of predict-observe-explain sequences in teams, while using laboratory materials such as magnets, compasses and nails (Botzer and Reiner 2005, p. 156). They were asked to describe the phenomena of magnetism verbally as well as to illustrate their explanations in sketches. During the predict sequence students were asked to develop a model that described what would happen in the different experimental conditions. In the comment field, students developed an explanatory model of the observed results by drawing schematic representations. Students' interactions were videotaped and all observations were complemented by field notes. Additionally, the authors collected students' descriptions and their sketches. Using these different data, the authors analyzed the number of categories representing students' understanding of magnetic phenomena, the relationship between visual representation and context as well as systematic patterns in the development of students' visual representations during the learning activity. Their results indicated that students' qualitative understanding can be supported by integrating different kinds of visual representations.

Mortimer and Buty (2009) analyzed learning difficulties while dealing with representations of the "infinite" in a teaching unit of ray optics. Representation of the infinite means that in an experimental setup for creating real images via a convex lens either the position of the image or the position of the object can approach the infinite (Mortimer and Buty 2009, p. 231). The authors videotaped episodes of 15 learning activities in a secondary II class, in which students created graphical representations of the course of light rays through a converging lens in teams. The authors considered the following learning mechanisms to be effective for explaining learning progress: first, an exchange between the world of theory and models and the world of concrete objects and events, second, switching between different forms of representation and, third, communicative exchanges which triggered learners to deal with different perspectives.

In a qualitative study of pupils' representations in the domain of particle models about solids, liquids, and gases, Waldrip et al. (2010) showed that pupil-generated representations can support conceptual learning if the teacher fosters the clarity, coherence and adequacy of pupils' concepts (Waldrip et al. 2010, p. 71 f.). Hubber et al. (2010, p. 23) confirmed the efficacy of using MRs in mechanics while teaching or learning the concept of force in a qualitative video study based on the observation of 12 lessons given by three teachers.

In conclusion, working with MRs explicitly fosters the development of conceptual understanding in different age groups and in various fields of physics. First, it is helpful to connect several representations that are either presented in different formats

(Mortimer and Buty 2009; Waldrip et al. 2010) or can be assigned to different levels of abstraction (Wilhelm 2005; Mortimer and Buty 2009). Second, it is helpful to encourage pupils to generate their own representations and to discuss them in the classroom (Botzer and Reiner 2005; Mortimer and Buty 2009; Waldrip et al. 2010; Hubber et al. 2010).

The strategies mentioned above can be understood as a contribution to cognitive activation (Lipowsky 2009, p. 93; Baumert and Kunter 2011, p. 13). Various mainly quantitative, quasi-experimental studies in the domain of mathematics have also provided clear evidence for the effectiveness of these strategies (Hiebert and Wearne 1993; Stein and Lane 1996; Shayer and Adhami 2007; Baumert and Kunter 2011).

However, in the domain of physics, most of the studies outlined above are based on small sample sizes and mainly qualitative analyses. For this reason, further research is needed to clarify the effectiveness of these strategies. This article attempts to help close this gap.

In this context, the present study compares conceptual understanding with an approach based on cognitive activation through MR learning tasks (referred to as treatment group TG A) with the conceptual understanding of learners working on learning tasks about the same content, dealing with the same variety and number of intuitive concepts during the similar time-on-task, but without the specific MR-based cognitive activation measures (treatment group B). Furthermore, we compare both treatment groups A) and B) with a control group C) from a related second study (study II), in which pupils learn with conventional tasks without addressing conceptual obstacles. In this setting we want to explore the following research questions:

1. Which kind of treatment A) or B) is best suited for fostering pupils' conceptual understanding in ray optics?
2. Are both kinds of treatment A) and B) more efficient for fostering pupils' conceptual understanding in ray optics than learning with the conventional tasks C) of study II?

The sample, instructional materials, and analysis techniques are described in the present section.

### 3 Methods

#### 3.1 Sample

Pupils in both studies were in their first year of physics lessons. This means that they were at the end of grade seven or at the beginning of grade eight depending on the school ( $M=13.45$  years,  $SD=0.67$ ). In both studies, the intervention was embedded in the regular curriculum and started when the topic of image formation by a convex lens was taught regularly in the school year. Pupils participated voluntarily in the data collection process, providing their parents had given their permission beforehand. To ensure that measurements could be related to each other, all data were pseudonymized.

**Table 1** Structure and size of the sample

	Level	<i>N</i>	<i>n</i> (study I)		<i>n</i> (study II)
		All	TG A) <sup>a</sup>	TG B) <sup>a</sup>	CG C) <sup>b</sup>
5	Schools	16	10	10	6
4	Teachers	18	10	10	8
3	Classes	29	11	10	8
2	Pupils	729	268	243	218
1	Measurements	1994	750	678	566

<sup>a</sup>Treatment Group<sup>b</sup>Control Group

The participating pupils were attending either German “Gymnasium”<sup>2</sup> (91 %) or comprehensive school<sup>3</sup> (9%). The number of girls was slightly greater than the number of boys ( $n_{girls}=371$ ,  $n_{boys}=358$ ). As the study was embedded in the regular curriculum, all school classes were assigned as a whole to the different conditions. For this reason, we have a multilevel structure in both studies: measurements are nested in pupils; pupils are nested in classes, and classes are nested in schools (see Table 1). In study I, each teacher taught at least one treatment class A) and one treatment class B). As we only took the control group of study II into account, we had eight different teachers in study II. Teachers’ professional experience varied between 2–30 years in study I and 4–25 years in study II. Higher and lower achieving classes were balanced in both studies based on the average level of former school grades per class before the intervention began. This approach was applied to ensure that pupils in both studies were equal in terms of knowledge and achievement when comparing the different groups within study I and study II. However, for study II, we only took into account the control group C) in this analysis.

### 3.2 Design and Instructional Material

For both studies, a quasi-experimental, repeated-measures design was used. The same version of the concept test (see 3.3) was applied three times. The measuring points were directly before and after the intervention (pre-post: short term), as well as 6–8 weeks later (follow-up: medium term). The covariates were the relevant school grades in math, physics, and German taken from the most recent school report. Furthermore, we investigated whether gender or class size had an influence on the development of pupils’ conceptual understanding.

Pupils in all conditions worked on tasks about forming real images using a convex lens during a sequence of six lessons (6 × 45 min) based on physics experiments. In study I, 63 % of the total intervention time was spent on learning tasks, which were different in TG A) and TG B). 37 % of the intervention time pupils did the same activities. During the same overall intervention time (6 × 45 min), pupils in CG C) worked on conventional tasks and carried out the same physics experiment as the pupils in study I. The teachers in all conditions implemented a detailed lesson plan, which the investigators handed out and discussed with the teachers before the intervention began.



**Table 2** Percentage of time spent on generating own representations, dealing with provided representations, or completing representations. 100% refers to the amount of time that TG A) and B) spent on different tasks (63% of the total intervention time)

	Self-generated	Provided	Completing
TG A)	58%	9%	33%
TG B)	14%	44%	42%

**Table 3** Kinds of representations per condition: 100% refers to the number of representations that play a role in the time spent on working on different tasks in TG A) and TG B) (63% of the total intervention time)

Kind of representation	TG A)	TG B)
Depictive (logic, schematic)	39%	33%
Depictive (realistic)	5%	4%
Descriptive (mathematical, numeric)	9%	9%
Descriptive (mathematical, symbolic)	7%	7%
Descriptive (verbal)	40%	47%

Study I involved a treatment condition TG A) in which pupils learnt with representational tasks focusing on widespread intuitive concepts, enhanced by four different measures of cognitive activation (explained below). Pupils in TG B) worked on learning tasks about the same content and dealt with the same variety and number of intuitive concepts, but without the specific cognitive activation measures. Altogether TG A) and B) both addressed seven intuitive concepts, which relate to (1) problems understanding the relation between light propagation and perception as well as scattering and perception, (2) problems understanding light propagation in terms of rays as a scientific model, e.g., confusing light rays (model) and light beams (phenomenon), (3) problems understanding the emergence of virtual images using a convex lens (4) as well as problems understanding the emergence of real images using a convex lens, such as the image passing the lens as a whole, (5) if part of the lens is covered, only half of the real image appears (6) or a real image cannot be formed without a screen (7). All intuitive concepts were taken from existing research in this domain (Goldberg and McDermott 1987, p. 111 f.; Wiesner 1992, p. 16 f.; Reiner et al. 2000, p. 14).

Study II provided a control group CG C) that did not work on conceptual difficulties at all (see Scheid 2013, p. 124 f.). CG C) worked on the same content but dealt with only one representational format at a time. There were two reasons why we took CG C) as a second comparison. First, we wanted to have a baseline that would enable us to see whether both learning strategies employed in TG A) and TG B) are helpful for pupils compared with conventional tasks used in classrooms. Second, we wanted to analyze how large possible differences in learning gains might be to determine whether it is worth using these kinds of tasks in classroom practice.

Tables 2 and 3 give an overview of the different representational task formats used in study I (TG A and TG B). Table 2 shows how often different representations in different formats were used; these frequencies were largely balanced between TG A) and TG B). The small differences result from the different learning activities required (see Table 3). There were three kinds of essentially different learning activities with representations: generating own representations, dealing with provided representations, or completing representations. Table 3 shows how often these different activity

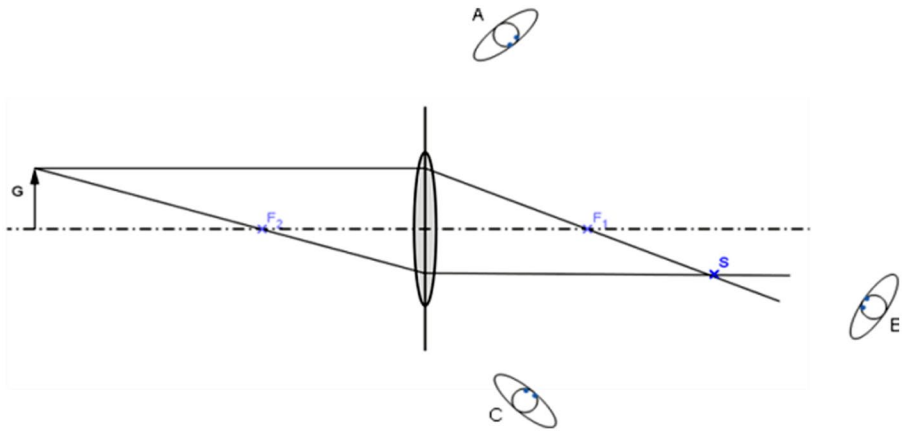
formats were required of pupils. Pupils in TG A) mainly worked with self-generated representations (58% of the time) and only 9% of the time with representations provided by the teacher, whereas TG B) worked more with representations provided by the teacher than their own representations (44% vs. 14%, respectively). The quantity of representations that had to be completed (e.g., to complete a ray diagram) was 33% in TG A) and 42% in TG B). All instructional material was prepared by the authors in the form of worksheets and overhead transparencies, and validated in a feedback loop with the participant teachers.

One main difference between TG A) and TG B) was the extent to which pupils generated representations themselves, which in fact was one of the cognitive activation strategies to enhance the representational learning activities. We now turn to a more detailed description of these strategies. As explained before, seven different learning difficulties related to common intuitive concepts were addressed; however, not all of the tasks dealt with an intuitive concept at the same time. In fact, this was the case for 56% of the tasks that differed between TG A) and TG B).

Type 1: Pupils in TG A) were explicitly asked to create their own representations and to reflect on them; Hiebert and Wearne (1993, p. 419), Stein and Lane (1996, p. 71) as well as Shayer and Adhmi (2007, p. 274 f.) reported on the effectiveness of this strategy in the field of mathematics. For example, in one task, pupils in TG A) were asked to create a sketch of the experimental setup used to create real images via a convex lens, whereas pupils in TG B) worked on a given sketch of this set-up and were asked to name the objects and the related optical quantities.

Type 2: For another strategy, pupils compared their own representations with a given solution, to reflect the own view and to adapt the own reasoning when necessary (see Hiebert and Wearne 1993; Baumert and Kunter 2011); these steps were supported by a whole-class discussion. For instance, in order to learn how to construct a ray diagram, TG A) was asked to complete a partly given solution. In the next step, they compared their own ray diagram with a worked out solution. In order to provide a scaffold for self-generation, pupils worked in three steps (with a different image size in each case) on a kind of representational fading-out task: First, 50% of the ray diagram was presented and 50% had to be completed; second, only 25% of the ray diagram was presented and 75% had to be completed; third, the ray diagram had to be constructed entirely. In each case, pupils went through the reflection process and compared their solution with the correctly worked out solution, which was handed out to the pupils afterwards. Pupils in TG B) were informed about how to construct a ray diagram first before they had to construct a ray diagram themselves. Furthermore, they measured all optical quantities, which is a conventional standard task. In another task of this type, the pupils dealt with the intuitive concept that only part of the image would appear if the lens was covered; see intuitive concept (6). Whereas pupils in TG A) constructed a ray diagram themselves, TG B) worked with a given ray diagram that showed that the entire image appeared on the screen.

Type 3: This cognitive activation strategy consisted of asking pupils to connect different kinds of representations to each other and to translate one representation into another (Hiebert and Wearne 1993; Stein and Lane 1996; Baumert and Kunter 2011). For example, in another task related to the example for type 2 (construction of a ray diagram), pupils had to complete a verbal formulation of one of the underlying

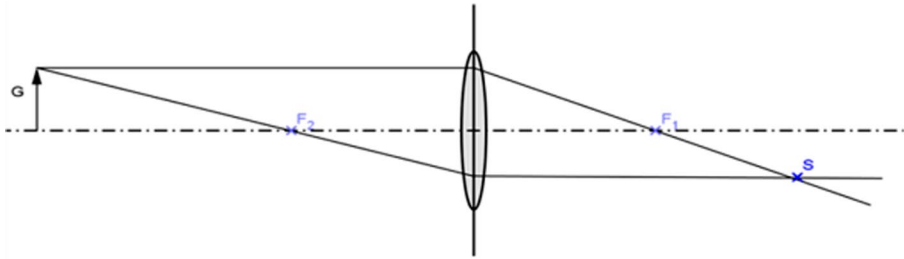


**Fig. 2** “Schematic representation of a real image formed by using a convex lens in TG A). ( $F$ =focal point,  $G$ =“Gegenstand” (object, here a candle)

rules for this construction (relating focal length as well as image, object sizes and distances). This process required them to relate the underlying experiment, the ray diagram and a descriptive representation as a third representational format in TG A).

Type 4: This type of task focused on dealing flexibly with a given representational format (Hiebert and Wearne 1993; Stein and Lane 1996; Baumert and Kunter 2011) and required pupils to create an internal mental model of the situation (Gentner and Gentner 1983, Mortimer and Buty 2009). For example, TG A) was asked which observer could see an image of a candle in an experimental setup when a candle (here represented by an arrow) was placed in front of a convex lens (see Fig. 2) (a) if someone put an opaque screen in position S, (b) if the opaque screen was replaced by a transparent one, and (c) if the screen was removed (referring to observers A, B, and C, respectively, in Fig. 2)<sup>4</sup>. Many pupils believed that the screen was necessary to “capture” the image, which is a widespread intuitive concept (Goldberg and McDermott 1987, p. 114). They did not consider that it would still be possible to see the so-called aerial image from an observer position where the light bundle emanating from the image points would hit the eyes. Moreover, TG A) was asked to work with the ray diagram in Fig. 2 in order to explain which of the observers (A–C) could see the optical image on the screen. Projecting themselves into the observer’s perspective required pupils to develop a mental model of the experimental situation to answer the question as to what the observer could see. At the same time, TG B) was asked where the image would form and from where it would be visible. Hence, the same conceptual difficulty, see intuitive concept 7), was addressed, but pupils were not required to operate on the ray diagram and the observers (A–C) were not shown in the provided representation (see Fig. 3).

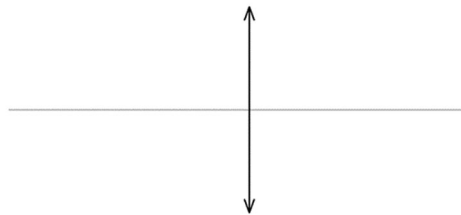
In another task of type 4, pupils in TG A) were asked to imagine what happens to the size and the position of a real image formed by convex lens, if the object, a candle, is moved towards to or away from the lens. This task required them to create a mental model of the situation for a “mental simulation”. Pupils in TG A) were told to display their reasoning graphically. In the next step, they were asked to imagine an



**Fig. 3** <sup>a</sup>Schematic representation of a real image formed by using a convex lens in TG B). (<sup>a</sup> $F$ = focal point,  $G$ =“Gegenstand” (object, here a candle)

**Fig. 4** <sup>a</sup> Conventional tasks in CG C) demanding the construction of a ray diagram (depictive logic, schematic representation). (<sup>a</sup> $f$ focal length,  $G$ object size,  $B$  image size,  $g$  distance between object and lens,  $b$  distance between image and lens)

Are you able to find out the size of the image without an experiment?  
 Draw a ray diagram!  
 $f = 3\text{ cm}$ ,  $g = 9\text{ cm}$ ,  $G = 2\text{ cm}$



**Fig. 5** <sup>a</sup>Conventional tasks in CG C) requiring pupils to deal with mathematical, descriptive, symbolic, and numeric representation(s). (<sup>a</sup> $f$ focal length,  $G$  object size,  $B$  image size,  $g$  distance between object and lens,  $b$  distance between image and lens)

Complete the table! Use the the equation for calculating the scale of magnification or reduction of a real image.

object size	image size	distance between object and lens	distance between image and lens
$G$	$B$	$g$	$b$
3 cm	6 cm		12 cm
3 cm		2 cm	5 cm
	2 cm	500 cm	4 cm

extreme case in which the candle was placed at a very far distance and to outline the effects on the image distance. At the same time, TG B) worked on two ray diagrams; they were asked to construct the image of a magnified candle light in a first setting (strengthening routine procedures) and of the focal point in another setting (a more challenging task). Thus, TG B) dealt with a problem-solving task as well and also worked on ray diagrams. However, there was no requirement to simulate changes and dependencies in a mental model.

During the identical intervention time, the pupils in CG C) were also exposed to the same representations (such as the photo of the experimental setup, ray diagrams or the related equation, etc.). Their lessons included the same experiment, but they neither worked simultaneously with more than one representation nor dealt with intuitive concepts. Instead, CG C) solved tasks related to image construction with principle rays (see Fig. 4) or calculated a missing quantity when three quantities were

known (see Fig. 5) (both conventional tasks). A detailed description of all tasks in CG C) can be found in Scheid (2013, p. 96 f.).

### 3.3 Analysis techniques

In order to assess pupils' conceptual understanding, a concept test that aimed at capturing conceptual knowledge in this domain was developed and used in both studies. Conceptual knowledge is defined as the knowledge of core concepts of a domain as well as an understanding of the mutual relationships among these core concepts (Byrnes and Wasik 1991, p. 777). The test was designed as a multiple-choice test and included 11 items, each of which had a scientifically correct answer and three distractors as answer options (see Table 4 for example items). These distractors were based on widespread intuitive concepts of pupils reported in the literature (Goldberg and McDermott 1987, p. 111 f.; Wiesner 1992, p. 16 f.; Reiner et al. 2000, p. 14). The testing time was 15 min.

To draw a conclusion from the comparison, we had to fulfill the following prerequisites:

1. We had to verify whether the concept test outlined above was a reliable and valid testing instrument for the evaluation of pupils' learning outcomes.
2. Pupils in all conditions had to start with the same level of prior knowledge.

In addition, when carrying out the comparison, the hierarchical structure of the data had to be taken into account: Measurement times were nested in pupils; pupils were nested in classes; classes were nested in schools. For this reason, the data were analyzed by multilevel analysis. All statistical calculations were conducted with the statistics software R, especially the package nlme (Pinheiro and Bates 2013). In the following analysis, only a 3-level model is considered, as the sample size of  $N=29$  on the class level is almost appropriate for multilevel modeling, but not the sample size on school level ( $N=18$ ); see Eid et al. (2011, p. 705).

The effect size  $\Delta$  described by Tymms (2004) was used to judge the relevance of significant effects. In analogy to Cohen's,  $\Delta=0.20$  is considered a small effect,  $\Delta=0.50$  a medium effect, and  $\Delta=0.80$  a large effect (see Bortz and Döring 2005, p. 568). According to Tymms (2004, p. 56 f.),  $\Delta$  can be calculated by dividing the difference of the requested means by the pooled standard deviation.

Missing data were handled as follows: (i) In terms of missing values on the level of individuals and measurement points, if all covariates were available, data of pupils, who were present for at least one measurement were taken into account. For instance, if all grades were known, but the pupil did not attend the pretest, data from the post- and follow-up tests could be included in the statistical model. For this reason, the sample size varied between the pre-, post-, and follow-up tests. (ii) In terms of missing values on the level of items in the concept test, unanswered questions were scored as 0 as long as pupils had attended the respective pre-, post-, or follow-up tests.

School grades and class size were grand-mean centered. After centering, the intercept can be interpreted as the expected value of the dependent variable of an average student who is in the group, in which all dummy coded categorical variables assume the value 0 (here sex=female, school type=comprehensive school, and condition

**Table 4** Items used in the concept test

Scientific concept	Item (example), all distractors were based on widespread intuitive concepts
Light propagation and ray model <sup>a</sup>	Which explanation is correct? <input type="checkbox"/> Light rays are something real, like thin water jets from a spray gun <input type="checkbox"/> Light rays exist only in people's minds, e.g., like constructions in geometry <input type="checkbox"/> Light rays are exactly the same as light beams <input type="checkbox"/> Light beams are mental objects; for example, they are used to determine the image size
Scattering and scientific explanation of visual perception <sup>b</sup>	Which of the following objects/creatures can be seen in a completely dark room? <input type="checkbox"/> A glowing firefly <input type="checkbox"/> A white sheet of paper <input type="checkbox"/> A bicycle reflector <input type="checkbox"/> The eyes of a cat
Real images formed via convex lens <sup>c</sup>	In an experimental assembly, a light bulb, a convex lens, and a screen are installed so that an enlarged, upside down, and sharp image of the filament is formed What will happen if the bottom half of the lens is covered? <input type="checkbox"/> The upper half will be cut off <input type="checkbox"/> The bottom half will be cut off <input type="checkbox"/> The image will become darker <input type="checkbox"/> The image will become brighter

<sup>a</sup>Item from scale (RD): understanding the model of light as rays and ray diagrams

<sup>b</sup>Item from scale (PS): understanding basic concepts of light propagation and scattering

<sup>c</sup>Item from scale (CL): understanding image formation while part of the lens is covered

=CG C). Another advantage of centering is that in multilevel models, the variance of intercepts can be interpreted (Hox 2010, p. 62).

## Results

To ensure that the first prerequisite was fulfilled, the concept test instrument was analyzed using data from all pupils in study I and study II ( $N=988$ ). The internal consistency as an estimator for reliability was satisfactory ( $\alpha_c=0.75$ ). A further validation analysis was carried out (e.g., cross-validation of explanatory and confirmatory factor analysis, see Hettmannsperger 2015, p. 207 f.), leading to three clearly interpretable dimensions: “understanding image formation while a part of the lens is covered” (referring to one of the most widespread intuitive concepts), “understanding basic concepts of light propagation and scattering” as well as “interpreting light according to the ray model and understanding ray diagrams” (see Table 5).

With regard to the second prerequisite, results in Fig. 6 show that pupils in the three conditions did not enter into the quasi-experimental activities with different levels of knowledge; we did not find any significant differences in the pretest.

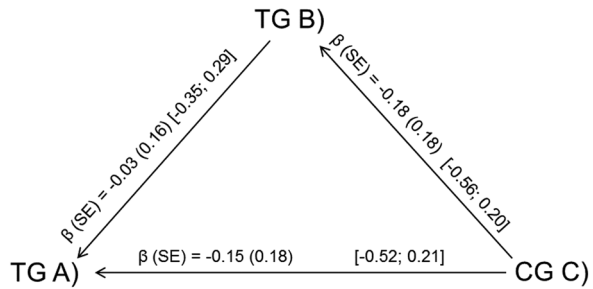
Figure 6 as well as Figs. 8 and 9 should be interpreted as follows. The direction of each arrow describes the difference between the group at the starting point (tail) of the arrow and the head of the arrow. Positive coefficients show that the group at the arrow head outperformed the group at the starting point of the arrow; this means that the difference from the starting to the end point (arrow head) has to be added. Nega-

**Table 5** Subscales derived from the cross validation of explanatory and confirmatory factor analyses

Factor	Interpretation	Items	$\alpha$
1 (CL)	Understanding image formation while part of the lens is covered	$x_{15a}, x_{15b}, x_{15c}$	0.81
2 (PS)	Understanding basic concepts of light propagation and scattering	$x_2, x_3, x_6, x_9, x_{10}$	0.65
3 (RD)	Understanding the model of light as rays and ray diagrams	$x_4, x_{11}, x_{14}$	0.49

**Fig. 6** Comparison between all three conditions TG A), TG B), and CG C) in the pretest, TG (treatment group), CG (control group)

differences at  $t_1$  (before the start of the intervention)



The format of the data presented in the triangle is as follows:  
 $\beta$  (Standard Error) [.] confidence interval of  $\beta$ : confidence level = 0.95  
 \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ ;  
 $\Delta$  (only reported if  $\beta$  significant)

tive coefficients indicate the opposite effect: The group at the arrow head improved less than the group at the starting point of the arrow.

In the first step of the actual analysis, the descriptive results (see Table 6) indicated that the pupils in both studies might have improved their conceptual understanding (main effect of time).

Descriptive results shown as box plots (see Fig. 7) indicate that pupils in both treatment groups in study I might have improved more than those in study II. Moreover, the descriptive results and the box plots suggest that this effect was stable with regard to the follow-up test.

In the second step, a multilevel analysis was performed to find out whether the trends in step 1 were significant. The values of the intra-class correlation ( $ICC_{Level-3} = 0.21$  and  $ICC_{Level-2} = 0.53$ ) indicated that a multilevel approach was indeed appropriate for analyzing the data (see Eid et al. 2011, p. 705).

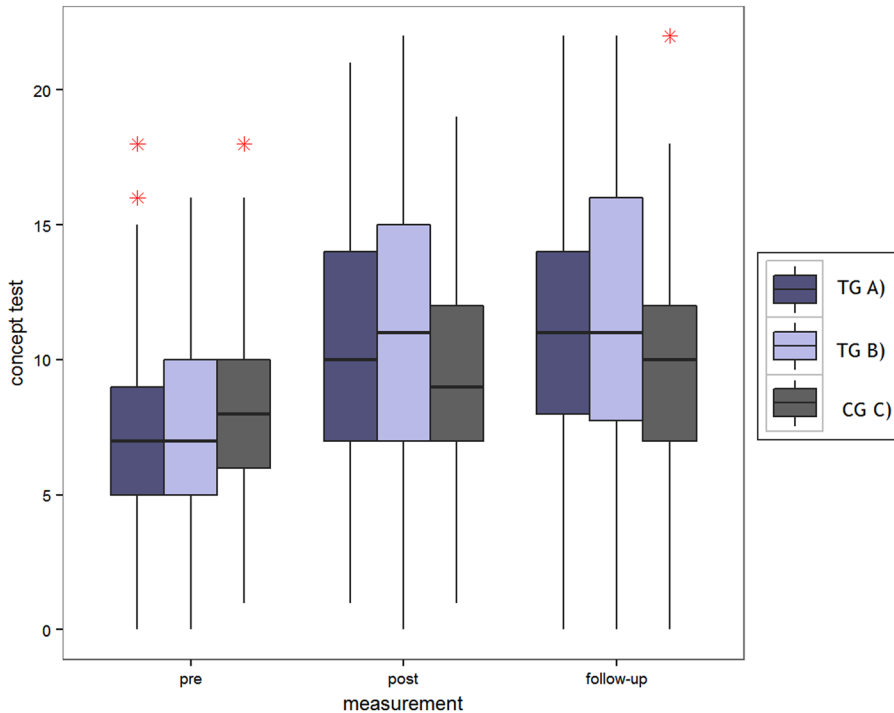
The results in Table 7 show that grades in mathematics and physics were significant covariates. German grades, class-size, and gender were not found to have any significant influence on conceptual understanding. Furthermore, there was a significant main effect of time: pupils improved their understanding overall in both the pre–post and the pre–follow-up comparison (see Table 7).

To find out which of the three conditions was best, we had to analyze the interaction between time and condition in step 3.

With regard to our first research question, which involved determining which kind of treatment was the most effective (addressing pupils’ intuitive concepts via repre-

**Table 6** Descriptive statistics of the concept test results

	Study I				Study II	
	TG A)		TG B)		CG C)	
	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>
Pre	7.47	(3.22)	7.42	(3.23)	8.02	(2.98)
Post	10.73	(4.70)	11.21	(5.33)	9.36	(3.63)
Follow-up	10.87	(4.65)	11.24	(4.86)	9.55	(3.37)

**Fig. 7** Box plots comparing the results of the concept test between all three conditions TG A), TG B), and CG C), TG (treatment group), CG (control group). (\*outliers)

sentational cognitive activation in TG A) versus addressing them with learning tasks without this instructional measure TG B)), we did not find any significant differences. Pupils in TG A) did not gain more or less knowledge pre—post ( $\beta = -0.10$ ,  $SE = 0.09$ ,  $F(1, 249) = 1.42$ ,  $p = 0.233$ ) and pre—follow-up ( $\beta = -0.07$ ,  $SE = 0.08$ ,  $F(1, 129) = 0.71$ ,  $p = 0.400$ ) than pupils in TG B) (see Fig. 8 and Fig. 9).

With regard to the second research question, the assumption that learning in TG A) and TG B) was more effective than learning with conventional tasks in CG C), we found significant effects in the expected direction. Pupils in TG A) improved more pre—post ( $\beta = 0.57$ ,  $SE = 0.10$ ,  $F(1, 1249) = 31.67$ ,  $p < 0.001$ ,  $\Delta = 0.68$ ) as well as pre—follow-up ( $\beta = 0.53$ ,  $SE = 0.10$ ,  $F(1, 1249) = 29.61$ ,  $p < 0.001$ ,  $\Delta = 0.63$ ) than CG C). The same was true for pupils in TG B) pre—post ( $\beta = 0.67$ ,  $SE = 0.10$ ,  $F(1, 1249) = 43.90$ ,  $p < 0.001$ ,  $\Delta = 0.80$ ) as well as pre—follow-up ( $\beta = 0.60$ ,  $SE = 0.10$ ,



**Table 7** Results of the Multilevel Analysis,  $\beta$  (standardized coefficient), SE (Standard Error)

	$t_1$ (pre) $\beta$ (SE)	CI <sup>a</sup>	$t_2$ (post) $\beta$ (SE)	CI <sup>a</sup>	$t_3$ (follow-up) $\beta$ (SE)	CI <sup>a</sup>
Intercept/main effect of time	1.62 (0.12) ***	[1.39; 1.85]	0.78 (0.07) ***	[0.64; 0.92]	0.90 (0.07) ***	[0.76; 1.04]
TG B) vs A)	-0.03 (0.16)	[-0.35; 0.29]	-0.10 (0.09)	[-0.27; 0.07]	-0.07 (0.08)	[-0.24; 0.09]
CG C) vs TG A)	-0.15 (0.18)	[-0.52; 0.21]	0.57 (0.10) ***	[0.37; 0.77]	(0.53) (0.10) ***	[0.34.; 0.72]
CG C) vs TG B)	-0.18 (0.18)	[-0.56; 0.20]	0.67 (0.09) ***	[0.47; 0.87]	0.60 (0.10) ***	[0.41; 0.79]
Maths	-0.09 (0.03) *	[-0.16; -0.02]	-0.02 (0.05)	[-0.11; 0.07]	-0.09 (0.05) *	[-0.19; 0.00]
Physics	-0.02 (0.03)	[-0.08; 0.05]	-0.18 (0.05) ***	[-0.27; -0.09]	-0.13 (0.04) ***	[-0.21; 0.05]
German	-0.02 (0.03)	[-0.08; 0.04]	-0.10 (0.05)	[-0.09; -0.01]	-0.05 (0.04)	[-0.13; 0.04]
Gender	0.04 (0.06)	[-0.07; 0.15]	0.15 (0.08)	[0.00; 0.30]	-0.04 (0.07)	[-0.19; 0.10]
Class size	0.05 (0.07)	[-0.09; 0.20]	-0.08 (0.04)	[0.00; 0.16]	-0.09 (0.04)	[0.01; 0.17]

Random effects (standardized):  $\sigma_u$  (class) = 0.33,  $\sigma_p$  (individual) = 0.41,  $\sigma_{\epsilon}$  (residual) = 0.50; Intra-class correlation (ICC)  $_{Level-3}$  = 0.21 and  $_{Level-2}$  = 0.53

\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$

<sup>a</sup>CI = confidence interval

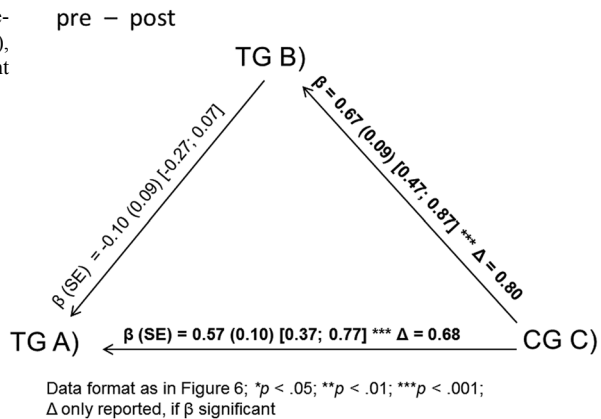
$F(1, 1249) = 36.86, p < 0.001, \Delta = 0.72$ ). The values of  $\Delta$  even suggest that the lessons in TG B) were more effective than in TG A). However, this effect was not significant (see Fig. 8 and Fig. 9).

In summary, the results in Fig. 8 indicate that with regard to the pre–post learning increase, the pupils in sample I (TG A and TG B) improved significantly more than pupils learning with conventional tasks (CG C). The same was true for the pre–follow-up comparison. In each case, we found a statistically highly significant increase of around half a standard deviation and a medium effect size  $\Delta$ . However, we did not find any significant difference between treatments A) and B) in study I. The box plots in Fig. 7 and the values of the coefficients  $\beta$  (see Fig. 8 and Fig. 9) display a small trend that the pupils in TG B) improved slightly more pre–post and pre–follow-up than the pupils in the TG A). As mentioned above, this effect was not significant.

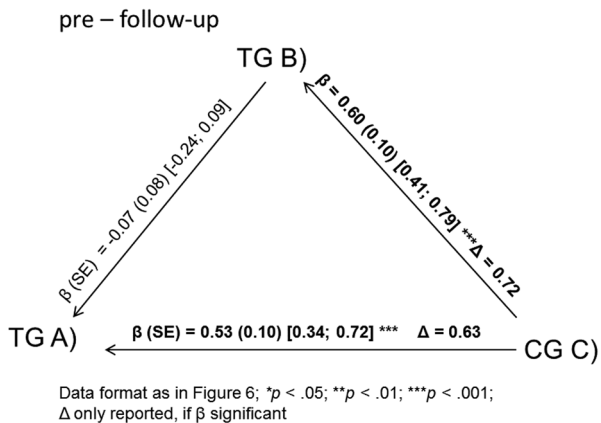
### Discussion

The comparison between TG A) and B) (research question 1) revealed that addressing conceptual difficulties via cognitive activation through learning with MRs is as effective as addressing them with learning tasks without this instructional measure. On the one hand, this is good news as multi-representational reasoning requires an additional cognitive activity and thus creates not additional cognitive load, which can be potentially harmful for learning. On the other hand, in view of the general potential of representational reasoning for science learning, one might have expected

**Fig. 8** Pre–post comparison between all three conditions TG A), TG B), and CG C), TG (treatment group), CG (control group)



**Fig. 9** Pre–follow-up comparison between all three conditions TG A), TG B), and CG C), TG (treatment group), CG (control group)



TG A) to attain a better conceptual understanding than TG B). There might be at least two possible explanations for this finding.

First, the difference between both kinds of treatment A) and B) might have been too “small” in terms of cognitive activation, as addressing pupils’ intuitive concepts can be seen as a kind of cognitive activation itself (Lipowsky 2009, p. 94; Baumert and Kunter 2011, p. 13).

Second, some of the tasks for TG B) required pupils to develop a mental model of the situation in order to answer the question. Developing a correct mental model of the situation implicitly also requires pupils to develop conceptual understanding and to deal with representations to a certain extent. The difference between TG A) and TG B) is therefore that representational reasoning on intuitive concepts is explicitly required in the former case, while it might be implicitly needed in the latter. Again this would lead to a too small difference between the treatments.

The comparison of both treatments A) and B) with the CG C) (research question 2) indicated that pupils learning with tasks addressing widespread intuitive concepts gained a better conceptual understanding than pupils in the control group. These differences were in each case highly significant and of practical relevance (medium

effect size). Moreover, they remained temporally stable beyond a short-term effect (6–8 weeks).

In summary, the question regarding the possible difference between the two treatment forms cannot be answered in the present study. However, we may conclude that addressing intuitive concepts with (multi-)representational reasoning was performed as an instructional measure for fostering conceptual learning in a usual physics classroom. First, the comparison to addressing the same intuitive concepts without explicit representational reasoning showed that the extra requirement did not constitute a harmful cognitive load, and second, the comparison pre – post (and pre – follow-up) showed reasonable practical relevance (medium effect size). Even though the finding that an explicit strategy improved conceptual understanding might be not very surprising, we have to keep in mind that overcoming widespread intuitive concepts is notoriously difficult and classical strategies such as inducing cognitive conflict by showing demonstration experiments do not automatically lead to success (Limón 2001; Vosniadou 2013). Compared to the range of results about conceptual change strategies known from meta-analysis ( $d \leq 0.6$ ; Hattie 2009), the result of the present approach turned out to be at the upper end. Thus, if an educational objective such as conceptual change is rather hard to achieve, also moderate steps towards improvement are welcome. As reasoning with MRs seems essential also for a series of other objectives in science education, it is a promising state of affairs that positive effects on conceptual understanding are among the benefits of this instructional approach and can possibly be combined with this kind of cognitive activation measures.

Future investigations could try to combine conceptual learning with other objectives of representational learning in science such as developing a coherent understanding of an experiment and the various representational formats related to it. Moreover, it would be interesting to compare the use of self-generated representations and representations provided by teachers in more detail. In view of the multiple functions of MRs, such attempts might be of interest for classroom practice and might help to clarify which kind of cognitive processes can be enhanced by which kind of representational learning activity.

## Notes

- 1 Learners believing in the impetus concept assumed that an object with a constant velocity inherits a property, e.g. a force, which maintains the velocity. Objects slow down by “using up” this property.
- 2 Secondary-school track for high achievers regularly leading to an A-level.
- 3 Secondary-school track first leading to the General Certificate of Secondary Education (GSCE) in grade 10 that offers qualified pupils to pass the A-Levels after another 3 years of schooling.
- 4 The observers A, B and C in Fig. 2 should not be mixed up with the quasi-experimental conditions TG A), TG B) and CG C).

**Acknowledgments** Support by Amanda Habbershaw and Rebecca Cors for proofreading the final English version of the manuscript is gratefully acknowledged.

## References

- Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, 33(2–3), 131–152.
- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183–198.
- Ainsworth, S. (2008). The educational value of multiple representations when learning complex scientific concepts. In J. Gilbert, M. Reiner, & M. Nakhleh (Eds.), *Visualization: Theory and practice in science education* (pp. 191–208). Dordrecht: Springer.
- Baumert, J., & Kunter, M. (2011). Das Kompetenzmodell von COACTIV [The COACTIV Competence Model]. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (pp. 29–53). Münster: Waxmann.
- Bortz, J., & Döring, N. (2005). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler [Research Methods and Evaluation for Human and Social Scientists]* (3rd edition). Heidelberg: Springer.
- Botzer, G., & Reiner, M. (2005). Imagery in physics learning—From physicists' practice to intuitive pupils' understanding. In J. K. Gilbert (Ed.), *Visualization in science education* (pp. 147–168). Dordrecht: Springer.
- Byrnes, J. P., & Wasik, B. A. (1991). Role of conceptual knowledge in mathematical procedural learning. *Developmental Psychology*, 27(5), 777–786.
- Cheng, M., & Gilbert, J. K. (2009). Towards a better utilization of diagrams in research into the use of representative levels in chemical education. In J. K. Gilbert & D. F. Treagust (Eds.), *Multiple representations in chemical education* (pp. 55–73). Dordrecht: Springer.
- Cheng, P. C.-H., & Shipstone, D. M. (2003). Supporting learning and promoting conceptual change with box and AVOW diagrams. Part 1: Representational design and instructional approaches. *International Journal of Science Education*, 25(2), 193–204.
- Cox, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction*, 9(4), 343–363.
- Dutke, S. (1994). *Mentale Modelle: Konstrukte des Wissens und Verstehens: kognitionspsychologische Grundlagen für die Software-Ergonomie [Mental Models. Constructs of knowledge and understanding]*. Göttingen: Verlag für Angewandte Psychologie.
- Eid, M., Gollwitzer, M., & Schmitt, M. (2011). *Statistik und Forschungsmethoden [Statistics and Research Methods]* (2nd edition). Weinheim: Beltz.
- Galili, I., & Hazan, A. (2000). Learners' knowledge in optics: Interpretation, structure and analysis. *International Journal of Science Education*, 22(1), 57–88.
- Gentner, D., & Gentner, D. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *In mental models* (pp. 99–129). Hillsdale: Lawrence Erlbaum Associates.
- Gilbert, J. K., & Treagust, D. (Eds.) (2009). *Multiple representations in chemical education*. Berlin: Springer.
- Goldberg, F. M., & McDermott, L. C. (1987). An investigation of student understanding of the real image formed by a converging lens or concave mirror. *American Journal of Physics*, 55, 108–119.
- Hattie, A. C. (2009). *Visible Learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hettmannsperger, R. (2015). *Lernen mit multiplen Repräsentationen aus Experimenten: Ein Beitrag zum Verstehen physikalischer Konzepte*. Wiesbaden: Springer VS.
- Hiebert, J., & Wearne, D. (1993). Instructional Tasks, Classroom Discourse, and Pupils' Learning in Second-Grade Arithmetic. *American Educational Research Journal*, 30(2), 393–425.
- Hox, J. J. (2010). *Multilevel analysis* (2nd edition). London: Routledge Academic.
- Hubber, P., Tytler, R., & Haslam, F. (2010). Teaching and Learning about Force with a Representational Focus: Pedagogy and Teacher Change. *Research in Science Education*, 40, 5–28.
- Kim, E., & Pak, S.-J. (2002). Students do not overcome conceptual difficulties after solving 1000 traditional problems. *American Journal of Physics*, 70(7), 759–765.
- Lee, G., Kwon, J., Park, S. S., Kim, J. W., Kwon, H. G., & Park, H. K. (2003). Development of an instrument for measuring cognitive conflict in secondary-level science classes. *Journal of Research in Science Teaching*, 40, 585–603.
- Leisen, J. (1998). Förderung des Sprachlernens durch den Wechsel von Symbolisierungsformen im Physikunterricht [Fostering language learning by changing the format of symbolisation in physics classrooms]. *Praxis der Naturwissenschaften Physik*, 47(2), 9–13.

- Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning and Instruction, 11*, 357–380.
- Lipowsky, F. (2009). Unterricht. In E. Wild (Ed.), *Pädagogische Psychologie [Educational Psychology]* (1st edition, pp. 73–102). Berlin: Springer.
- Mortimer, E. F., & Buty, C. (2009). What does “In the Infinite” Mean? The difficulties with dealing with the representation of the “Infinite” in a teaching sequence on optics. In C. Andersen, N. Scheuer, M. D. P. Pérez Echeverría, & E. Teubal (Eds.), *Representational systems and practices as learning tools* (pp. 225–242). Rotterdam: Sense Publishers.
- Özdemir, G., & Clark, D. B. (2007). An overview of conceptual change theories. *Eurasia Journal of Mathematics, Science & Technology Education, 3*(4), 351–361.
- Pinheiro, J., & Bates, D. M. (2013). nlme [Computer Software]: The R Project for Statistical Computing. <http://cran.r-project.org/web/packages/nlme/nlme.pdf>. Accessed 17. Sept 2011.
- Plötzner, R., & Spada, H. (1998). Inhalt, Struktur und Anwendung von Physikwissen: Eine psychologische Perspektive [Content, Structure and Application of Physics Knowledge: A Psychological Perspective]. *Zeitschrift für Didaktik der Naturwissenschaften, 4*(2), 81–100.
- Reiner, M., Slotta, J. D., Chi, M. T. H., & Resnick, L. B. (2000). Intuitive physics reasoning: A commitment to substance-Based conceptions. *Cognition and Instruction, 18*, 1–34.
- Scheid, J. (2013). Multiple Repräsentationen, Verständnis physikalischer Experimente und kognitive Aktivierung: Ein Beitrag zur Entwicklung der Aufgabenkultur. In H. Niedderer, H. Fischler, & E. Sumfleht (Eds.), *Studien zum Physik- und Chemielernen* (Vol. 151). Berlin: Logos Verlag.
- Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 49–69). Cambridge: Cambridge University Press.
- Schnotz, W. (2006). Conceptual change. In D. H. Rost (Ed.), *Handwörterbuch pädagogische Psychologie* (3rd edition, pp. 75–81). Weinheim: Beltz.
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*(2), 141–156.
- Sell, K. S., Herbert, B. E., Stuessy, C. L., & Schielack, J. (2006). Supporting student conceptual model development of complex earth system through the use of multiple representations and inquiry. *Journal of Geoscience Education, 54*(3), 396–407.
- Shayer, M., & Adhami, M. (2007). Fostering cognitive development through the context of mathematics: Results of the came project. *Educational Studies in Mathematics, 64*(3), 265–291.
- Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation, 2*(1), 50–80.
- Taber, K. (2009). Learning at the symbolic level. In J. K. Gilbert & D. F. Treagust (Eds.), *Multiple representations in chemical education* (pp. 75–105). Dordrecht: Springer.
- Thagard, P. (1991). Concepts and conceptual change (reprint of 1990 paper). In J. Fetzer (Ed.), *Epistemology and cognition* (pp. 101–120). Dordrecht: Kluwer.
- Tsui, C., & Treagust, D. (2013). *Multiple representations in biological education*. Dordrecht: Springer.
- Tymms, P. (2004). Effect sizes in multilevel models. In I. Schagen & K. Elliot (Eds.), *But what does it mean? The use of effect sizes in educational research* (pp. 55–66). University of London: Institute of Education.
- Vosniadou, S. (2013). Reframing the classical approach to conceptual change: Preconceptions, misconceptions and synthetic models. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education*. Berlin: Springer.
- Waldrup, B., Prain, V., & Carolan, J. (2010). Using multi-modal representations to improve learning in junior secondary science. *Research in Science Education, 40*, 65–80.
- Wiesner, H. (1992). Schülervorstellungen und Lernschwierigkeiten mit dem Spiegelbild [Pupil Preconceptions and Learning Difficulties Regarding Mirror Imaging]. *Naturwissenschaften im Unterricht—Physik, 3*, 16–18.
- Wilhelm, T. (2005). *Konzeption und Evaluation eines Kinematik/Dynamik-Lehrgangs zur Veränderung von Schülervorstellungen mit Hilfe dynamisch ikonischer Repräsentationen und graphischer Modellbildung* [Conceptualization and Evaluation of a Kinematics/Dynamics Course for Changing Pupil Preconceptions with the Help of Dynamically Iconic Representations and Graphical Model Construction] (Doctoral dissertation). Julius-Maximilians-Universität Würzburg, Germany.
- Zimrot, R., & Ashkenazi, G. (2007). Interactive lecture demonstrations: A tool for exploring and enhancing conceptual change. *Chemistry Education Research and Practice, 8*(2), 197–211.