



When learning from animations is more successful than learning from static pictures: learning the specifics of change

Rolf Ploetzner¹ · Sandra Berney² · Mireille Bétrancourt²

Received: 4 September 2020 / Accepted: 4 May 2021 / Published online: 26 May 2021
© The Author(s) 2021

Abstract

The results of three meta-analyses show that the effectiveness of learning from animations, when compared to learning from static pictures, is rather limited. A recent re-analysis of one of these meta-analyses, however, supports that learning from animations is considerably more effective than learning from static pictures if the specifics of the displayed changes need to be learned. In order to further validate this finding as well as to clarify the educational strengths and weaknesses of animations and static pictures, an experimental study with three groups was conducted. Overall, 88 university students participated in the study. One group of learners ($n=30$) watched a single picture of a gear mechanism, one group of learners ($n=28$) watched four pictures, and one group of learners ($n=30$) watched an animation. All groups had to identify specific motions and spatial arrangements covered by the gear mechanism. While learners who watched the animation exhibited the best performance with respect to the identification of motions, learners who watched the pictures showed the best performance with respect to the identification of spatial arrangements. The effect sizes are large. The results of the study help to clarify when animations and when static pictures are most suitable for learning.

Keywords Learning · Animations · Static pictures · Specifics of change · Educational effectiveness

✉ Rolf Ploetzner
rolf.ploetzner@ph-freiburg.de

Sandra Berney
sandra.berney@unige.ch

Mireille Bétrancourt
mireille.betrancourt@unige.ch

¹ Department of Psychology, University of Education, Kunzenweg 21, 79117 Freiburg, Germany

² TECFA-FAPSE, University of Geneva, 40, Boulevard du Pont d'Arve, 1211 Geneva, Switzerland

Introduction

In education, animations are employed to support the perception, mental representation, and understanding of changes in space and time. However, the educational effectiveness of animations has been challenged by three meta-analyses which compared learning from animations with learning from static pictures. In the first meta-analysis, Höffler and Leutner (2007) analyzed 76 pair-wise comparisons. They observed a merely small overall effect size of $d=0.37$ in favor of animations and videos. Almost 10 years later, the second analysis by Berney and Bétrancourt (2016) covered 140 pair-wise comparisons. It resulted in an even smaller overall effect size of $g=0.226$ in favor of animations. The third analysis by Castro-Alonso et al. (2019) involved 82 pair-wise comparisons. It also yielded an overall effect size of merely $g=0.23$ in favor of animations.

The disillusioning results of these meta-analyses resulted in an increasingly growing skepticism over the educational benefits of animations. For instance, Clark and Mayer (2016, p. 84) suggest "... to use static illustrations unless there is a compelling instructional rationale for animation. In particular, when you have an explanative illustration, we recommend presenting a series of static frames to depict the various states of the system rather than a lock-step animation." However, what denotes a compelling instructional rationale for the use of animation? One important reason for the educational use of animations is that learners need to grasp what animations can overtly present: how change in space and time occurs. Already 2002 Tversky, Morrison, and Bétrancourt assumed "... if there are benefits to animation, they should be evident especially for continuous rather than discrete changes, in particular, for manner of change and for microsteps, the subtle and intricate timing relations among parts of a complex system." (Tversky et al., 2002, p. 258). Until recently, however, the specific potential of animations has not received sufficient attention in empirical research comparing the educational effectiveness of animations and static pictures.

More recently, Ploetzner et al. (2020) re-analyzed the meta-analysis originally published by Berney and Bétrancourt (2016). They investigated a new moderator that encodes whether the specific features of the displayed changes were irrelevant or relevant to learning. They distinguished between simple and complex features with respect to the features of change. A frequent type of change is motion, for instance. If it merely had to be learned in which direction an object moves, this was coded as a simple feature of change relevant to learning. However, if it had to be learned whether an object moves slow or fast, or whether an object speeds up or slows down, this was coded as a complex feature of change relevant to learning.

Learning from animations was significantly more successful than learning from static pictures if either simple ($g=0.340$) or complex ($g=0.647$) features of the displayed changes had to be learned. If neither simple nor complex features of the presented changes had to be learned, it was irrelevant whether learning took place on the basis of animations or static pictures ($g=0.043$). These results suggest that when the learning domain includes only simple forms of change, many learners who are presented static pictures seem to be able to construct suitable mental animations. However, when challenging forms of change are to be learned, mental animation is more likely to be difficult and prone to error, thus learning from animated displays is more beneficial (cf. Hegarty et al., 2003).

The authors conclude that their re-analysis is of heuristic value and that future research on learning from animations needs to validate this finding by means of experimental studies. In this paper, we present an experimental study that compares learning from an

animation with learning from static pictures. It aims to experimentally challenge the findings reported by Ploetzner et al. (2020) as well as to investigate the potential and limitations of animations for learning in relation to static pictures. The study is therefore not only based on the representational characteristics of animations and static pictures, but also on a model of how animations are perceptually and cognitively processed.

In the following sections, the theoretical background is described. Thereafter, the study that experimentally compares learning from an animation with learning from static pictures is presented. A discussion and conclusions complete the paper.

Theoretical background

Animations and static pictures have different representational characteristics. Both display visuospatial information (cf. Ploetzner & Lowe, 2012; Schnotz & Lowe, 2008). Visuospatial information refers to the set of graphic entities that make up the display as well as how these entities are arranged in space. Spatial arrangements can be specified by referring to the absolute positions of entities in the display (e.g. an entity is located in the upper left corner) or by referring to the relative positions of entities (e.g. an entity is located left of another entity). Furthermore, visual spatial information refers to how entities are spatially organized (e.g. an entity is made up of two other entities).

In contrast to static pictures, animations display not only visuospatial but also spatiotemporal information, i.e., changes in space over time (cf. Scaife & Rogers, 1996). Animations consist of a sequence of static pictures in which each picture differs slightly from the preceding picture. If the separate pictures are displayed at a sufficient rate (e.g. 24 pictures per second), humans experience the optical illusion of continuous change in the display. Spatiotemporal information refers to the set of events that constitute an animation. Events denote entities and how they change over time. During events, entities might change (e.g. an entity gets larger) and/or the spatial arrangements of entities might change (e.g. an entity rotates around another entity). Furthermore, spatiotemporal information refers to how events are temporarily organized (e.g. an event takes place before another event or two events take place simultaneously).

Whereas animations display changes explicitly, static pictures can merely indicate them by means of arrows or onion skins, for example (cf. Jenkinson, 2017). Hence, when learning from animations, the learners can directly perceive the changes. In contrast, learners have to infer the changes when learning from static pictures – a process that is often demanding and prone to error (cf. Hegarty et al., 2003). Thus, animations are more informative than static pictures with regard to the changes that occur in the displayed subject matter (see also Kühl et al., 2018). Therefore, it is frequently assumed that learning from animations is more effective than learning from static pictures if the displayed subject matter involves changes in space and time.

Even if an animation explicitly depicts visuospatial and spatiotemporal information, in order to be educationally effective, it needs to be sufficiently perceived and comprehended. The Animation Process Model (APM; Lowe & Boucheix, 2008, 2011; Lowe & Schnotz, 2014) describes how learning from unnarrated animations progresses in order to construct a more and more complete mental model of the presented subject matter (see also Kriz & Hegarty, 2007). The model describes five phases in which perceptual bottom-up and cognitive top-down processes interact. However, it is not assumed that

learners process these phases in linear order. Especially learners with only little prior knowledge will have to repeatedly apply the different processes before adequate understanding is reached.

According to Lowe and Boucheix (2008, 2011), during Phase 1 the learners identify confined event units which may be presented at different spatial and temporal locations. Event units represent graphic entities and the behavior they exhibit. If the learners possess only little prior knowledge about the displayed subject matter, the separation of event units will mostly be a bottom-up process. That is, it relies mainly on the perceptual properties of the visual display, such as the colors and sizes of presented areas or the relative rates at which areas in the display change.

The event units identified in Phase 1 make up the basic components for the succeeding phase. In Phase 2, event units are gradually and iteratively merged into larger but still restricted structures. Essential to this activity is the construction of visuospatial and spatiotemporal relations that depend on the perceptual properties of the animated display. For instance, event units that are close to each other in space or time may be combined into spatiotemporal structures named dynamic micro-chunks. During Phase 3, spatially and temporarily distributed dynamic micro-chunks are combined to produce more extensive relational structures such as causal chains. This demands the use of domain-relevant general knowledge. The iterative combination of established relational structures can finally embrace the animation's entire spatial and temporal extent, yielding a global characterization of the animation.

In Phase 4, by taking advantage of domain-specific prior knowledge, the learners associate functional tasks with the established structures. As a consequence, the structures are described as functional episodes which represent the functionality of the displayed subject matter. During Phase 5, the learners refine the established functionality to identify the conditions under which the animated system operates, for example. This may result in a mental model of the animated subject matter that is complete, coherent, and consistent. Subsequently, such a model can be applied to new but analogous situations.

In studies of learning from animation, the participants are mostly beginners with respect to the animated subject area. In order to compensate for the learners' lack of domain-specific knowledge, the dynamic visualizations are often combined with spoken or written narrations. The narrations may guide the learners' attention to specific events in the display, comment and describe events, or explain relationships between events such as causes and effects. If an animation involves both pictorial and verbal information, the learners need to sufficiently relate both sources of information. The cognitive theory of multimedia learning (Mayer, 2009, 2014) and the integrated model of text and picture comprehension (Schnotz, 2014) both delineate the perceptual processes as well as the cognitive processes that are important to learning from pictorial and verbal information.

The distinction between perceptual and cognitive processes drawn in the APM emphasizes that learning at the perceptual level and learning at the cognitive level can make up educational objectives in their own rights. Although education very often strives for the acquisition of conceptual models, the meta-analyses by Höffler and Leutner (2007), Berney and Bétrancourt (2016), and Castro-Alonso et al. (2019), together with the meta-analysis by Ploetzner et al. (2020) indicate that the teaching of conceptual models is not a specific strength of animations. Instead, the meta-analysis by Ploetzner et al. (2020) suggests that animations can be more successfully employed if kinematic models need to be learned, i.e. if the learners need to construct mental representations of the displayed changes and how they unfold in time (cf. Hegarty & Just, 1993; Hegarty et al., 2003).

Experimental study

Research questions and hypotheses

Due to the different representational characteristics of animations and static pictures, the educational effectiveness of animations and static pictures might depend on the learning tasks that have to be accomplished. The meta-analysis conducted by Ploetzner et al. (2020) gives raise to the assumption that learning from animations is more successful than learning from static pictures if the specifics of the displayed changes have to be learned. It is hypothesized that this result is due to the fact that animations display spatiotemporal information completely and explicitly. Static pictures, in contrast, depict spatiotemporal information merely incompletely and implicitly.

Both animations and static pictures display visuospatial information completely and explicitly. It can, therefore, be hypothesized that visuospatial arrangements can be learned equally well from animations and static pictures. It might even be that visuospatial information is more successfully learned from static pictures than from animations. Visuospatial arrangements are stationary in static pictures. As a consequence, the learners can attend to and identify them without being visually distracted. In contrast, visuospatial arrangements are permanently changing in animations. These changes might make it difficult for the learners to attend to and identify specific arrangements at specific points in time. That is, while animations might lead the learners to focus on spatiotemporal aspects, static pictures might lead the learners to focus on visuospatial aspects (cf. Ploetzner & Fillisch, 2017).

Thus, two hypotheses were tested in an experimental study:

- (1) If spatiotemporal information – especially the specifics of the displayed change – are relevant to the learning task, then learning from an animation is expected to be more successful than learning from static pictures.
- (2) If visuospatial information – such as spatial arrangements – has to be learned, then learning from static pictures is expected to be equally or more successful than learning from an animation.

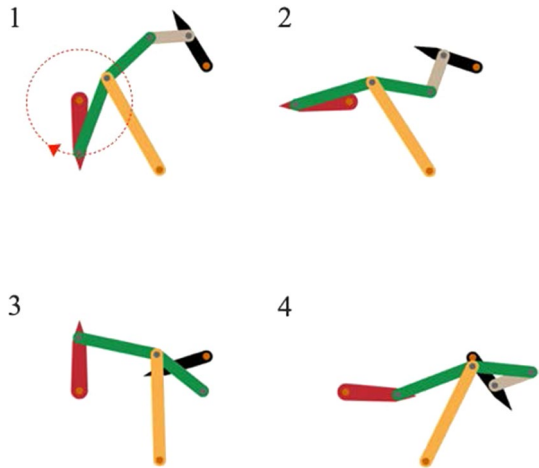
Learning task

The learning task refers to a six-bar linkage, a gear mechanism that moves in a plane (cf. Fig. 1). Six-bar linkages are constructed from six links – including the rack – and seven joints (cf. Volmer, 1992). They very often transform continuous rotation into complex motion.

The employed six-bar linkage¹ converts continuous clockwise rotation of the input, or drive gear (red link), into discontinuous counterclockwise rotation of the output gear (black link) with three halts of different durations. The transformation is realized by two non-uniformly moving couplers (green link and beige link) and an asymmetrically oscillating crank (yellow link). The links are connected by three joints fixed to the rack (orange joints) and four joints moving in a plane (grey joints).

¹ A video of the original linkage can be watched in the Digital Mechanism and Gear Library (<https://www.dmg-lib.org/dmglib/handler?manim=209022>).

Fig. 1 A sequence of four states of the six-bar linkage employed in the study (reconstructed with permission of the digital mechanism and gear library, www.dmg-lib.org)



Thus, with the exception of the input gear, the links move non-uniformly, asymmetrically, or even discontinuously. In machine engineering, gear atlases were traditionally used to describe the motion of linkages by static diagrams (e.g. Hain, 1972). These diagrams depict the paths of links as well as the relative gradients of links over time, for example. However, because even machine engineers may have difficulty in inferring the motion of linkages on the basis of static representations (Brix et al., 2005), digital gear libraries have been established in order to dynamically present the linkages (e.g. www.dmg-lib.org).

The learning task consisted of watching either one picture of the six-bar linkage, four pictures, or an animation. Thereafter, the learners had to identify the correct motion of each link out of four motions. Furthermore, the learners had to identify the correct arrangement of each pair of joined links out of four arrangements.

Pre-study

In the main study, the learners have to identify the motion of each link out of four different motions as well as the arrangement of each pair of joined links out of four different arrangements. It was therefore investigated in a pre-study, as to whether the learners could sufficiently distinguish between the different motions and arrangements they were shown.

Design

Two groups of learners were investigated. While one group had to distinguish between different motions of each link, the other group had to distinguish between different arrangements of each pair of joined links.

Participants

A total of 33 students volunteered for the study and received financial compensation for their participation. All students were enrolled in undergraduate pre-service teacher programs in the STEM and other disciplines at a university in southwest Germany. The students were randomly assigned to the group that had to distinguish between motions (14

females and 3 males, mean age $M=21.29$ years, $SD=3.33$) and the group that had to distinguish between arrangements (13 females and 3 males, mean age $M=23.25$ years, $SD=2.86$).

Material

The six-bar linkage consists of five moving links. With respect to each link four kinds of motion were animated: (1) continuous uniform motion (red, green, grey, and black link) or continuous symmetrical oscillation (yellow link), (2) continuous non-uniform motion (red, green, grey, and black link) or continuous asymmetrical oscillation (yellow link), (3) mirror-inverted continuous, nonuniform motion (red, green, grey, and black link) or mirror-inverted, continuous asymmetrical oscillation (yellow link), and (4) discontinuous motion (all links). All animations of a link moved in the same direction and covered the same path. This resulted in $5 \times 4 = 20$ animations covering five correct and 15 incorrect motions. Each animation lasted 6.2 s.

Because the six-bar linkage consists of $n=5$ moving but joined links, $n-1=4$ pairs of joined links can be considered. With respect to each pair four kinds of arrangement were displayed: (1) correct joint and correct relative position, (2) correct joint and mirror-inverted relative position of one link, (3) incorrect joint and correct relative position, and (4) incorrect joint and mirror-inverted relative position of one link. This resulted in $4 \times 4 = 16$ pictures covering four correct and 12 incorrect arrangements.

All animations of links and pictures of arrangements were produced with Adobe Animate CC and Adobe Illustrator CC. The left-hand side of Fig. 2 shows snapshots of the four animations of the green link at three seconds. Because the animations exhibit different motions, they only display the same positions and orientations at the start. The right-hand side of Fig. 2 shows the four arrangements of the green and yellow links. Of interest is if the learners are presented any of the four animations or any of the four arrangements, would they be able to distinguish the animation or arrangement they had watched from the other three animations or arrangements?

The learning tasks were presented to the learners by a computer program made with MatchWare Mediator 9. The size of the presentation area was 1200×1000 pixels. Initially, the use of the program and the learning task were described to the learners. Thereafter, an example of an animation or of an arrangement was shown for 31 s, i.e. the duration of five iterations of an animation. The examples relied on a four-bar linkage. Next, four animations or four arrangements were presented in quadrants labelled A, B, C, and D (cf. Fig. 3). In order to not confuse the learners by presenting four animations simultaneously, none of the animations or arrangements were initially visible. When the mouse was moved over a quadrant, the looping animation or the stationary arrangement became immediately visible. Thus, the learners were able to watch one animation or arrangement at a time. Furthermore, the learners were able to move the mouse back and forth between quadrants as long as they wished. After the learners decided which animation or arrangement corresponded to the one they had seen before, they received feedback as to whether their decision was correct.

After the example was finished, the 20 animations or the 16 arrangements described above were presented to the learners in random order. Each animation or arrangement was presented in the same manner as the animation or arrangement in the example. Only when all of the animations or arrangements were processed by the learners did they

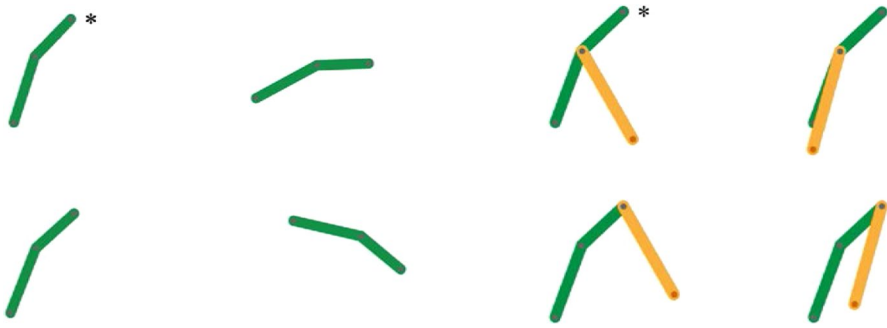


Fig. 2 Snapshots of four animations at three seconds (left) and pictures of four arrangements (right). The correct animation and arrangement are marked with an asterisk

receive feedback as to how many of their decisions were correct. Every correct response was scored with one point. The maximum score with respect to the identification of motions was 20 points; the maximum score with respect to the identification of arrangements was 16 points.

Procedure

Students participated in groups of up to 12 individuals. Each student was individually seated in front of a computer with a 21-inch screen. The computers were placed on separate tables. One group of students watched a randomized sequence of 20 moving links. After watching a moving link, the students had to identify the motion out of four motions. The other group of students watched a randomized sequence of 16 arrangements of pairs of joined links. After watching a pair of joined links, the students had to identify the arrangement out of four arrangements. The procedure took about 30 min.

Results

On average, 18.53 moving links (92.65%, $SD=1.18$) and 15.56 arrangements of pairs of joined links were identified correctly (97.25%, $SD=0.81$). The results demonstrate that the learners were able to sufficiently distinguish between the motions as well as the arrangements they were shown.

Main study

Design

Three groups of learners were investigated (cf. Fig. 4). In group ‘Picture’, the learners watched a single picture of the six-bar linkage. In group ‘Four Pictures’, the learners viewed a sequence of four pictures. In group ‘Animation’, the learners watched an animation.

To compare learning from an animation to learning from a single picture might be sound from a methodological point of view (cf. Castro-Alonso et al., 2016). However, it might be problematic from a psychological point of view. If an animation is replaced

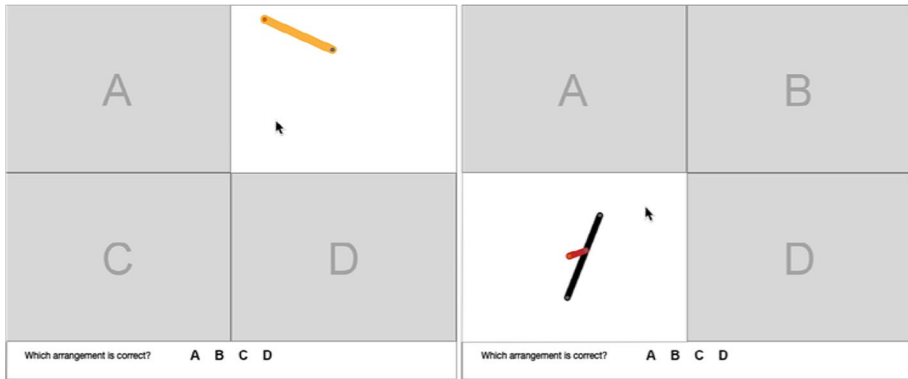


Fig. 3 The four quadrants with an example animation (left) and an example arrangement (right) uncovered (translated by the authors)

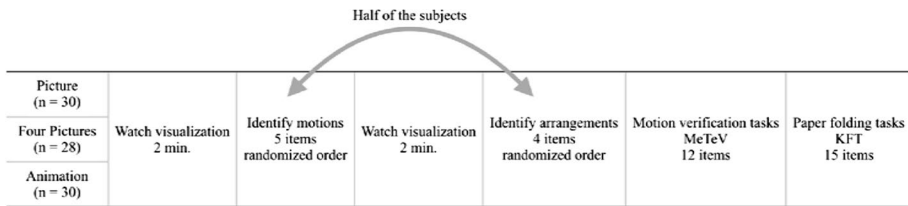


Fig. 4 The design of the main study

by pictures, information gets lost. The fewer pictures are presented, the more information gets lost. Novice learners are hardly able to infer motions from a single picture. Therefore, we also presented a sequence of four pictures of the six-bar linkage. This gave the learners – at least in principle – the opportunity to infer motions by comparing and contrasting the presented pictures (cf. Ploetzner & Lowe, 2014).

In each group, the learners had to accomplish two learning tasks and were therefore required to watch the static or dynamic visualization of the six-bar linkage twice. After one viewing, the learners had to identify the correct motion of each link. After the other viewing, the learners had to identify the correct arrangement of each pair of joined links. To counterbalance possible sequencing effects (e.g. Jhangiani et al., 2019), within each group, half of the learners received the learning tasks in inverse order than the other half. Furthermore, the learners’ mechanical ability and spatial ability were assessed as potential covariates.

Participants

A total of 88 students volunteered for the study and received financial compensation for their participation. All students were enrolled in undergraduate pre-service teacher programs in the STEM and other disciplines at a university in southwest Germany. The students were randomly assigned to the group ‘Picture’ (23 females and 7 males, mean

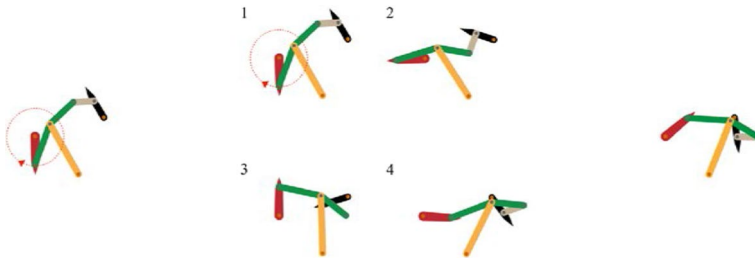


Fig. 5 How the picture (left), the sequence of four pictures (middle), and the animation (right) of the six-bar linkage were presented to the learners

age $M=21.47$ years, $SD=2.54$), the group ‘Four Pictures’ (21 females and 7 males, mean age $M=22.04$ years, $SD=2.37$), and the group ‘Animation’ (23 females and 7 males, mean age $M=21.63$ years, $SD=3.01$).

Material

Learning tasks The single picture, the sequence of four pictures, and the non-interactive animation of the six-bar linkage were produced with Adobe Illustrator CC and Adobe Animate CC. Figure 5 shows how the single picture (left-hand side), the sequence of four pictures (middle) and the animation (right-hand side) were presented to the learners. All visualizations of the six-bar linkage were of the same size. The single picture equals the first picture of the sequence of four pictures. The four pictures equal four frames shown in the animation. The single picture as well as the first picture of the sequence of pictures indicate the direction of rotation of the drive gear (red link) by a dashed circle with an arrow. The four pictures were numbered according to the succession of frames they depict. In all groups, the learners were additionally instructed in written form that the drive gear (red link) uniformly rotates in clockwise direction.

The five correct and 15 incorrect animations of individual links as well as the four correct and 12 incorrect pictures of arrangements of pairs of joined links were the same as those employed in the pre-study.

The learning tasks were presented to the learners by a computer program made with MatchWare Mediator 9. The size of the presentation area was 1200×1000 pixels. Initially, the use of the program and the learning tasks were described to the learners. Thereafter, a single picture, a sequence of four pictures, or an animation of an example four-bar linkage were presented to the learners for 90s. Next, the learners had to identify the motion of each link out of four motions or the arrangement of each pair of joined links out of four arrangements. All animations of individual links and all pictures of arrangements of pairs of joined links were presented in the same way as in the pre-study (cf. Fig. 3). After all animations or arrangements were processed by the learners, they received feedback as to how many of their decisions were correct.

After the example was finished, the single picture, the sequence of four pictures, or the animation of the six-bar linkage were shown to the learners for two minutes. Thereafter, the learners had to identify the motion of each link out of four motions or the arrangement of each pair of joined links out of four arrangements. Again, all animations of individual links and all pictures of arrangements of pairs of joined links were

presented in the same way as in the pre-study (cf. Fig. 3). The correct and incorrect animations of links and the correct and incorrect pictures of arrangements were presented in random order. After all animations or arrangements were processed by the learners, they received feedback as to how many of their decisions were correct. Every correct response was scored with one point. The maximum score with respect to the animated links was 5 points, the maximum score with respect to the identification of arrangements was 4 points.

Mechanical ability With respect to gear mechanisms, the learners' mechanical ability makes up a domain-relevant competence. This competence might influence how the learners imagine and perceive the motion of the six-bar linkage. Thus, the learners' mechanical ability was assessed by 12 motion verification tasks from the Test of Mechanical and Technical Understanding (MeTeV; Hartweg, 2010). Motion verification tasks can be solved on the basis of general mechanical principles (cf. Hegarty, 1992). Typically, they present a schematic picture of a mechanical system to the learners. The picture includes graphical or verbal information as to how certain components of the system move. The learners' task is to infer from the picture how other components of the system move. Hegarty et al. (1988) termed the inference processes required by motion verification tasks "mechanical reasoning" (see also Hegarty, 1992, 2004).

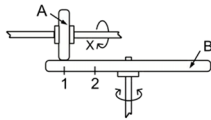
Figure 6 shows an example of a motion verification task taken from the Test of Mechanical and Technical Understanding. The format of all tasks was multiple-choice with four response options. Each task had to be processed within 90 s (cf. Hartweg, 2010). Every correct response was scored with one point. The maximum score was 12 points.

The Test of Mechanical and Technical Understanding has a reliability of 0.69 (Cronbach's Alpha, cf. Hartweg, 2010). It correlates 0.56 with spatial-visual ability as measured by the Mannheimer Test for the Assessment of Physical-Technical Problem Solving (Conrad et al., 1980) and 0.67 with mechanical-technical understanding as measured by the Wilde Intelligence Test 2 (Kersting et al., 2008).

Spatial ability Empirical research has demonstrated that learners with low or high spatial ability learn differently from static and dynamic visualizations (e.g. Höffler, 2010; Höffler & Leutner, 2011). Learners with high spatial ability are often more successful in constructing mental animations than learners with low spatial ability. As a consequence, the former learn more successfully from static pictures than the latter. In the present study, the learners' spatial ability was assessed by the Subtest N3 of the Cognitive Ability Test (KFT; Heller & Perleth, 2000). It employs 15 paper folding tasks comparable to those originally proposed by Ekstrom et al. (1976). The format of all tasks was multiple-choice with five response options. Each task had to be processed within 40 s (cf. Heller & Perleth, 2000). Every correct response was scored with one point. The maximum score was 15 points.

The Subtest N3 of the Cognitive Ability Test has a reliability of 0.79 (Kuder-Richardson, cf. Heller & Perleth, 2000). Because the validation of the initial version of the Cognitive Ability Test took place before the Subtest N3 was included, the test manual does not report the construct validity for the Subtest N3 (cf. Heller & Perleth, 2000).

The motion verification tasks and the paper folding tasks were presented to the learners by the same computer program that displayed the learning tasks. After each set of tasks, the learners received feedback as to how many tasks they solved correctly.



If Wheel A is shifted from Position 1 to Position 2, how does Wheel B move?

Wheel B rotates in a different direction.

Wheel B rotates slower.

Wheel B rotates faster.

Wheel B rotates as fast as before.

Question 8 of 12

Remaining time

Next

Fig. 6 An example motion verification task from the Test of Mechanical and Technical Understanding (MeTeV; reconstructed with permission of Verena Hartweg; translated by the authors)

Procedure

Students participated in groups of up to 16 individuals. Each student was individually seated in front of a computer with a 21-inch screen. The computers were placed on separate tables. The computers presented the material to the students in the following order: (1) first presentation of the visualization (one picture, four pictures, or animation), (2) identification of motions (first half of the students) or arrangements (second half of the students), (3) second presentation of the same visualization, (4) identification of motions (second half of the students) or arrangements (first half of the students), (5) motion verification tasks, and (6) paper folding tasks. The procedure took about 50 min.

Results

The means and standard deviations for each dependent variable are shown in Table 1. The intercorrelations among the dependent variables are provided in Table 2. While the group 'Animation' exhibited the best performance with respect to the identification of motions, the group 'Picture' showed the best performance with respect to the identification of arrangements.

The investigated groups did not significantly differ with respect to their mechanical ability ($F(2, 85)=0.73, p=0.484$) or with respect to their spatial ability ($F(2, 85)=1.16, p=0.319$). The correlations between the learners' mechanical ability and their identifications of motions as well as between the learners' mechanical ability and their identifications of arrangements were nonsignificant (cf. Table 2). Likewise, the correlations between the learners' spatial ability and their identifications of motions as well as between the learners' spatial ability and their identifications of arrangements were nonsignificant (cf.

Table 1 Means *M* and standard deviations *SD* of correct answers in each group

Variable	Picture		Four pictures		Animation	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Mechanical ability	5.27 (43.91%)	2.35	4.61 (38.42%)	2.06	4.73 (39.42%)	2.23
Spatial ability	7.70 (51.33%)	2.55	6.64 (44.27%)	2.56	6.83 (45.53%)	3.33
Identif. of motions	1.77 (35.40%)	0.94	1.64 (32.80%)	0.99	3.80 (76.00%)	0.80
Identif. of arrangements	3.93 (98.25%)	0.25	2.93 (73.25%)	1.27	2.93 (73.25%)	1.06

Table 2 Intercorrelations among the dependent variables

	Spatial ability	Identif. of motions	Identif. of arrangements
Mechanical ability	0.456**	0.087	0.184
Spatial ability		- 0.044	0.186
Identif. of motions			- 0.082

** $p < 0.01$

Table 2). Therefore, the learners' mechanical and spatial ability were not considered as covariates in the analysis of variance.

The results of a multivariate analysis of variance are shown in Table 3. On the multivariate level, the investigated groups differ significantly. The effect size is large (cf. Cohen, 1988; Rosenthal, 1994). On the univariate level, the groups differ significantly with respect to the identification of motions as well as with respect to the identification of arrangements. The effect sizes are large (cf. Cohen, 1988; Rosenthal, 1994). Concerning the identification of motions, post-hoc analyses further reveal that the group 'Animation' significantly differs from the group 'Picture' (Fisher's Least Significant Difference $LSD = 2.03$, $p < 0.01$, $d = 2.33$) as well as from the group 'Four Pictures' ($LSD = 2.16$, $p < 0.01$, $d = 2.41$). The group 'Picture' and the group 'Four Pictures' do not significantly differ from each other ($LSD = 0.12$, $p = 0.606$).

With regard to the identification of arrangements, post-hoc analyses reveal that the group 'Picture' significantly differs from the group 'Four Pictures' ($LSD = 1.00$, $p < 0.01$, $d = 1.11$) and the group 'Animation' ($LSD = 1.00$, $p < 0.01$, $d = 1.30$). The group 'Four Pictures' and the group 'Animation' do not differ from each other ($LSD = 0.00$, $p = 0.985$).

Discussion

The research reported in this paper originated from the fact that three meta-analyses which compared learning from animation to learning from static pictures found only small overall effect sizes in favor of animations (Berney & Bétrancourt, 2016; Castro-Alonso et al., 2019; Höffler & Leutner, 2007). At the same time, the production of animations can be much more challenging, time-consuming, and costly than the production of static graphics. The potentially harmful consequences that researchers as well as practitioners may draw from the above findings is that the production of animations is not worth the extra effort (cf. Clark & Mayer, 2016).

Table 3 Results of a multivariate analysis of variance (MANOVA)

Source of variation	Dependent variable	Wilks lambda	<i>df</i>	<i>F</i>	<i>p</i>	η^2	Cohen's <i>d</i>
Group	Multivariate	0.35	4.168	28.84	0.000	0.407	1.657
	Univariate	Identif. of motions	2.85	52.29	0.000	0.552	2.220
		Identif. of arrangements	2.85	11.17	0.000	0.208	1.025

In contrast, a recent meta-analysis by Ploetzner et al. (2020) found that learning from animation is considerably more effective than learning from static pictures if especially challenging features of change have to be learned. An experimental study was conducted in order to validate this finding. The learning task was made up of a mechanical device – a six-bar linkage – that produces accelerated, asymmetric, nonuniform, and discontinuous patterns of motion. Due to their irregularities, these patterns are difficult to predict by novices and even engineers (cf. Brix et al. 2005) on the basis of static pictures and everyday perceptual schemata (cf. Lowe & Schnotz, 2014). It was, therefore, hypothesized that watching an animation is more effective than watching static pictures for mentally representing the spatiotemporal information of the gear mechanism. However, because both animations and static pictures display visuospatial information completely and explicitly, it was hypothesized that watching an animation is not more effective than watching static pictures for mentally representing the visuospatial information of the gear mechanism. The results of the experimental study support both hypotheses and are in accord with the theoretical assumptions by Tversky et al. (2002) as well as with the results of the meta-analysis reported by Ploetzner et al. (2020).

While visuospatial information was more successfully learned from one picture than from the animation, learning from four pictures was not more successful than learning from the animation. This finding might indicate that both the four pictures and the animation provided abundant information to the learners. As a consequence, in both groups the learners might have had comparable difficulties in deciding on which information to focus and which information to extract.

The animation process model (APM, Lowe & Boucheix, 2008, 2011; Lowe & Schnotz, 2014) distinguishes between perceptual and cognitive processes during learning from animation. Especially, the APM emphasizes that learning at the perceptual level and learning at the cognitive level each make up educational objectives in their own rights. To demonstrate the potential of animations for learning at the perceptual level, the focus of the experimental study was on perceptual processing during learning from animation. Correspondingly, the employed tasks for identifying motions and arrangements have a certain proximity to perceptual memory tasks (e.g. Castro-Alonso & Atit, 2019; Schurgin, 2018). Conceptual processing, in contrast, was not investigated. Whether the neglect of conceptual processing is considered a limitation of the present study depends on the educational objectives. In many domains, perceptual processing and learning is an educational objective in its own right: e.g. motion patterns of technical devices in science and engineering (cf. Lowe & Boucheix, 2011; Ploetzner & Fillisch, 2017) as well as motion patterns of body parts in biology, medicine, physiotherapy, and physical training (cf. Chen & Wu, 2016; Sukel et al., 2003).

In many learning contexts, however, both perceptual and conceptual processing is required (e.g. how a technical device needs to be constructed in order to produce a certain motion pattern). Although the animation processing model (Lowe & Boucheix, 2008, 2011; Lowe & Schnotz, 2014) as well as the cognitive theory of multimedia learning (Mayer, 2009, 2014) abstractly address the interplay between perceptual and conceptual processing, it still remains to be described in more detail as to precisely how perceptual processes during learning from animations can facilitate conceptual understanding and vice versa. Therefore, future studies could not only focus on perceptual processing during learning from animation but on both perceptual and conceptual processing. Such studies could help in better understanding of how perceptual processing contributes to conceptual understanding.

Unexpectedly, the learners' spatial ability was not significantly related to learning performance. Especially with respect to the group 'Picture' and the group 'Four Pictures', it was expected that learners with higher spatial ability are more successful in mentally animating the relevant motions than learners with lower spatial ability (cf. Höffler, 2010; Höffler & Leutner, 2011). The descriptive data, together with the sporadic feedback of individual learners, give rise to the assumption that the learners in these groups were mostly overburdened with having to mentally animate the relevant motions. Even high spatial ability did not possibly help them in constructing suitable mental animations. Hence, this might indicate that even well-developed cognitive abilities cannot always compensate for learning material that is inadequate with respect to the learning task. Alternatively, because spatial ability constitutes a multidimensional construct (e.g. Buckley et al., 2018; Castro-Alonso & Atit, 2019), this might indicate that the employed assessment of spatial ability did not correspond well to the demands of the investigated learning situation.

The finding that instructional animations are especially beneficial when the specifics of change need to be learned certainly requires further validation. For instance, future studies could investigate learners with different characteristics (e.g. learners who possess more pre-knowledge about the investigated subject matter) or the learning of different subject matters. Furthermore, future studies could examine how learners construct mental representations of the observed movements. For instance, some learners might have tried to visually memorize the shown movements. Others might have tried to run a mental animation in order to mentally reconstruct the shown movements. Others might have produced internal verbalizations of the shown movements (e.g. the yellow link swings forth and back like a pendulum, fast from the right to the left and slow from the left to the right; cf. Lloyd-Jones et al., 2008). Still others might have used gestures (e.g. mimicking movements with their fingers) in order to support the encoding of the shown movements (cf. Brucker et al., 2015; Lajevardi et al., 2017). Learners might have taken advantage of such strategies either in isolation or in combination. Understanding better how learners construct mental representations while they process an animation perceptually could help in developing measures that support learners during this phase of learning.

After years of growing skepticism with respect to the educational effectiveness of animations, the reported findings may help to cast instructional animations in a more positive light again. Reliable empirical evidence about the instructional strengths of animations as well as of static pictures would certainly help multimedia designers and educators to decide when to make use of animations and when to resort to static pictures.

Conclusions

In past research, animations have been frequently employed in a way that their potential for learning was not fully realized. In particular, the representational characteristics of the animated display were often not well aligned with the demands of the learning task. As a consequence, learning from animations was found to be only modestly more effective than learning from static pictures. In accord with a meta-analysis conducted by Ploetzner et al. (2020), the experimental study presented in this paper demonstrates that learning from animation is significantly more successful than learning from static pictures if the specific features of the displayed change are to be learned. Thus, animations can be highly effective tools to support dynamic perceptual learning and the acquisition of kinematic models. Static pictures, in contrast, are especially promising in teaching spatial configurations.

Acknowledgements We thank Olesya Lang for supporting the data collection.

Author contributions All authors contributed to the conception of the study. The experimental design, the preparation of material, the data collection, the data analysis, and the writing of the first draft was performed by the first author. All authors contributed to the refinement of the draft and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability The data set is available from the Open Science Foundation: <https://doi.org/10.17605/OSF.IO/QP6UG>.

Code availability The text provides a link to the animation that underlies our reconstruction of the six-bar linkage. We have permission to reconstruct the linkage, to use the reconstruction in our research, and to include pictures of the reconstruction in publications. However, we have no permission to make the reconstruction itself available to the public. The employed mechanical ability and cognitive ability tests are available through the references provided in the text.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Berney, S., & Bétrancourt, M. (2016). Does animation enhance learning? A meta-analysis. *Computers & Education*, *101*, 150–167. <https://doi.org/10.1016/j.compedu.2016.06.005>
- Brix, T., Döring, U., & Trott, S. (2005). DMG-Lib: Ein moderner Wissensraum für die Getriebetechnik [DMG-Lib: A modern knowledge space for gear technology]. In Zentralbibliothek des Forschungszentrums Jülich (Hg.), Knowledge eXtended: Die Kooperation von Wissenschaftlern, Bibliothekaren und IT-Spezialisten [Knowledge eXtended: The cooperation of scientists, librarians, and specialists for information technology] (S. 251–262). Jülich: Forschungszentrum Jülich

- Brucker, B., Ehlis, A.-C., Häußinger, F. B., Fallgatter, A. J., & Gerjets, P. (2015). Watching corresponding gestures facilitates learning with animations by activating human mirror-neurons: An fNIRS study. *Learning and Instruction, 36*, 27–37. <https://doi.org/10.1016/j.learninstruc.2014.11.003>
- Buckley, J., Seery, N., & Cauty, D. (2018). A heuristic framework of spatial ability: A review and synthesis of spatial factor literature to support its translation into STEM education. *Educational Psychology Review, 30*, 947–972. <https://doi.org/10.1007/s10648-018-9432-z>
- Castro-Alonso, J. C., & Atit, K. (2019). Different abilities controlled by visuospatial processing. In J. C. Castro-Alonso (Ed.), *Visuospatial processing for education in health and natural sciences*. (pp. 23–51). Springer.
- Castro-Alonso, J. C., Ayres, P., & Paas, F. (2016). Comparing apples and oranges? A critical look at research on learning from statics versus animations. *Computers & Education, 102*, 234–243. <https://doi.org/10.1016/j.compedu.2016.09.004>
- Castro-Alonso, J. C., Wong, M., Adesope, O. O., Ayres, P., & Paas, F. (2019). Gender imbalance in instructional dynamic versus static visualizations: A meta-analysis. *Educational Psychology Review, 31*, 361–387. <https://doi.org/10.1007/s10648-019-09469-1>
- Chen, Y.-J., & Wu, J.-L. (2016). A computer-aided articulation learning system for subjects with articulation disorders. *Engineering Computations, 33*, 2185–2197. <https://doi.org/10.1108/EC-08-2015-0235>
- Clark, R. C., & Mayer, R. E. (2016). *E-learning and the science of instruction*. (4th ed.). Wiley.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. (2nd ed.). Lawrence Erlbaum Associates.
- Conrad, W., Baumann, E., & Mohr, V. (1980). *Mannheimer test zur erfassung des physikalisch-technischen problemlösens (MTP) [mannheimer test for the assessment of physical-technical problem solving]*. Hogrefe Verlag.
- Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Kit of factor-referenced cognitive tests*. Educational Testing Service.
- Hain, K. (1973). *Getriebebeispiel-atlas: Eine zusammenstellung ungleichförmig übersetzender getriebe für den konstruktur [atlas of gear examples: a collection of nonuniformly transmitting gears for the constructing engineer]*. VDI Verlag.
- Hartweg, V. (2010). Mechanisch-technisches Verständnis als Konstrukt in der testbasierten Studienberatung [Mechanical-technical ability as a construct in the test-based student advisory service] (Doctoral Dissertation). Aachen: Rheinisch-Westfälische Technische Hochschule Aachen.
- Hegarty, M. (1992). Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology, 18*, 1084–1102. <https://doi.org/10.1037/0278-7393.18.5.1084>
- Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences, 8*, 280–285. <https://doi.org/10.1016/j.tics.2004.04.001>
- Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language, 32*, 717–742. <https://doi.org/10.1006/jmla.1993.1036>
- Hegarty, M., Just, M. A., & Morrison, J. (1988). Mental models of mechanical systems: Individual differences in qualitative and quantitative reasoning. *Cognitive Psychology, 20*, 191–236. [https://doi.org/10.1016/0010-0285\(88\)90019-9](https://doi.org/10.1016/0010-0285(88)90019-9)
- Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction, 21*, 325–360. https://doi.org/10.1207/s1532690xci2104_1
- Heller, K. A., & Perleth, C. (2000). Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision [Cognitive ability test for grades 4 to 12, revision]. Göttingen: Beltz Test GmbH
- Höffler, T. N. (2010). Spatial ability: Its influence on learning with visualizations – A meta-analytic review. *Educational Psychology Review, 22*, 245–269. <https://doi.org/10.1007/s10648-010-9126-7>
- Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17*, 722–738. <https://doi.org/10.1016/j.learninstruc.2007.09.013>
- Höffler, T. N., & Leutner, D. (2011). The role of spatial ability in learning from instructional animations – Evidence for an ability-as-compensator hypothesis. *Computers in Human Behavior, 27*, 209–216. <https://doi.org/10.1016/j.chb.2010.07.042>
- Jenkinson, J. (2017). The role of craft-based knowledge in the design of dynamic visualizations. In R. Lowe & R. Ploetzner (Eds.), *Learning from dynamic visualization – Innovations in research and application*. (pp. 93–117). Springer.
- Jhangani, R. S., Chiang, I.-C.A., Cuttler, C., & Leighton, D. C. (2019). *Research methods in psychology*. (4th ed.). Kwantlen Polytechnic University.
- Kersting, M., Althoff, K., & Jäger, A. O. (2008). *Wilde-intelligenz-test 2 (WIT-2) [wilde intelligence test 2]*. Hogrefe Verlag.

- Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human Computer Studies*, *65*, 911–930. <https://doi.org/10.1016/j.ijhcs.2007.06.005>
- Kühl, T., Navratil, S. D., & Münzer, S. (2018). Animations and static pictures: The influence of prompting and time of testing. *Learning and Instruction*, *58*, 201–209. <https://doi.org/10.1016/j.learninstruc.2018.07.006>
- Lajevardi, N., Narang, N. S., Marcus, N., & Ayres, P. (2017). Can mimicking gestures facilitate learning from instructional animations and static graphics? *Computers & Education*, *110*, 64–76. <https://doi.org/10.1016/j.compedu.2017.03.010>
- Lloyd-Jones, T. J., Brandimonte, M. A., & Bäuml, K.-H. (2008). Verbalising visual memories. *European Journal of Cognitive Psychology*, *20*, 387–395. <https://doi.org/10.1080/09541440701755543>
- Lowe, R., & Boucheix, J.-M. (2008). Learning from animated diagrams: How are mental models built? In G. Stapleton, J. Howse, & J. Lee (Eds.), *Diagrammatic representations and inference – Proceedings of the fifth international conference, diagrams 2008*. (pp. 266–281). Springer.
- Lowe, R., & Boucheix, J.-M. (2011). Cueing complex animations: Does direction of attention foster learning processes? *Learning and Instruction*, *21*, 650–663. <https://doi.org/10.1016/j.learninstruc.2011.02.002>
- Lowe, R. K., & Schnotz, W. (2014). Animation principles in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning*. (2nd ed., pp. 513–546). Cambridge University Press.
- Mayer, R. E. (2009). *Multimedia learning*. (2nd ed.). Cambridge University Press.
- Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning*. (2nd ed., pp. 43–71). Cambridge University Press.
- Ploetzner, R., Berney, S., & Bétrancourt, M. (2020). A review of learning demands in instructional animations: The educational effectiveness of animations unfolds if the features of change need to be learned. *Journal of Computer Assisted Learning*, *36*, 838–860. <https://doi.org/10.1111/jcal.12476>
- Ploetzner, R., & Fillisch, B. (2017). Not the silver bullet: Learner-generated drawings make it difficult to understand broader spatiotemporal structures in complex animations. *Learning and Instruction*, *47*, 13–24. <https://doi.org/10.1016/j.learninstruc.2016.10.002>
- Ploetzner, R., & Lowe, R. (2012). A systematic characterisation of expository animations. *Computers in Human Behavior*, *28*, 781–794. <https://doi.org/10.1016/j.chb.2011.12.001>
- Ploetzner, R., & Lowe, R. (2014). Simultaneously presented animations facilitate the learning of higher-order relationships. *Computers in Human Behavior*, *34*, 12–22. <https://doi.org/10.1016/j.chb.2014.01.039>
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis*. (pp. 231–244). Russell Sage Foundation.
- Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies*, *45*, 185–213. <https://doi.org/10.1006/ijhc.1996.0048>
- Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning*. (2nd ed., pp. 72–103). Cambridge University Press.
- Schnotz, W., & Lowe, R. K. (2008). A unified view of learning from animated and static graphics. In R. K. Lowe & W. Schnotz (Eds.), *Learning with animation: Research implications for design*. (pp. 304–356). Cambridge University Press.
- Schurgin, M. W. (2018). Visual memory, the long and the short of it: A review of visual working memory and long-term memory. *Attention, Perception, & Psychophysics*, *80*, 1035–1056. <https://doi.org/10.3758/s13414-018-1522-y>
- Sukel, K. E., Catrambone, R., Essa, I., & Brostow, G. (2003). Presenting movement in a computer-based dance tutor. *International Journal of Human-Computer Interaction*, *15*, 433–452
- Tversky, B., Morrison, J., & Bétrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies*, *57*, 247–262. <https://doi.org/10.1006/ijhc.2002.1017>
- Volmer, J. (Ed.). (1992). *Getriebetechnik – Grundlagen [gear technology – basic principles]*. Verlag Technik.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.