Learning and Instruction 17 (2007) 722–738

Learning and Instruction

www.elsevier.com/locate/learninstruc

Instructional animation versus static pictures: A meta-analysis

Tim N. Höffler*, Detlev Leutner

nwu-Essen, Research Group and Graduate School on Science Education, Duisburg-Essen University, P.O. Box, D-45117 Essen, Germany

Abstract

A meta-analysis of 26 primary studies, yielding 76 pair-wise comparisons of dynamic and static visualizations, reveals a medium-sized overall advantage of instructional animations over static pictures. The mean weighted effect size on learning outcome is $d = 0.37$ (95% CI 0.25–0.49). Moderator analyses indicate even more substantial effect sizes when the animation is representational rather than decorational ($d = 0.40$, 95% CI 0.26–0.53), when the animation is highly realistic, e.g., video-based ($d = 0.76$, 95% CI 0.39–1.13), and/or when procedural-motor knowledge is to be acquired ($d = 1.06$, 95% CI 0.72–1.40). The results are in line with contemporary theories of cognitive load and multimedia learning, and they have practical implications for instructional design.

Keywords: Meta-analysis; Instructional visualization, animation, video, and picture; Declarative, procedural, and problem-solving knowledge

1. Introduction

In recent years there has been a lengthy debate about the opportunities for using animation in learning and instruction. The enthusiasm of the first years, in which the potential of dynamic visualization seemed to be boundless, gave way to a more pragmatic view. In particular, the review of Tversky, Morrison, and Bétrancourt (2002) influenced the instructional designers' community. The authors showed that animations often had no advantages over still pictures; but if they had, it was because more information was available in the animated than in the static version. Due to this result, the focus turned to the question of when dynamic displays are more effective in learning than static ones (Hegarty, 2004).

While there are many promising approaches to identifying such conditions (e.g., Ainsworth & VanLabeke, 2004; ChanLin, 2001; Lowe, 1999), there has been no systematic collection of these research results as yet. The present meta-analysis of 76 pair-wise comparisons of static pictures versus animations attempts to identify the factors responsible for successful learning with animations. In addition, it presents a comprehensive survey of studies comparing these two forms of visualization, and it analyzes which form may be superior in learning outcomes under what conditions.

* Corresponding author.
E-mail address: tim.hoeffler@uni-duisburg-essen.de (T.N. Höffler).

0959-4752/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.learninstruc.2007.09.013

2. Theoretical framework

2.1. The effect of pictures on learning outcome

Even though the research on learning with pictures has been conducted from different theoretical perspectives (see Levie & Lentz, 1982; Levin, Anglin, & Carney, 1987; Lewalter, 1997; Weidenmann, 2002), recent cognitive theories like Mayer's "Cognitive Theory of Multimedia Learning" (Mayer, 2001, 2005) or Schnotz's "Integrative Model of Text and Picture Comprehension" (Schnotz, 2005) can be used to describe and explain the results of a large number of studies. Mayer's theory, for example, regards the learner as a constructor of his or her own knowledge, actively selecting, organizing, and integrating relevant visual and verbal information. It is based on three basic assumptions:

- *Active processing*: According to Wittrock's (1974, 1989) generative theory of meaningful learning, learning occurs when learners actively process information (select – organize – integrate).
- *Dual channel processing and dual coding*: From Paivio's dual coding theory (Clark & Paivio, 1991; Paivio, 1986) and Baddeley's working-memory model (Baddeley, 1992), the notion of two different cognitive systems for information processing is taken: a verbal system transmitting and processing sequential information like written or spoken text and a visual system responsible for spatial information and images.
- *Limited capacity*: The overall information processing capacity is very strictly constrained by the limitations of short-term memory load within each system (Baddeley, 1992; Chandler & Sweller, 1991).

There is strong empirical evidence that learning outcomes are improved by presenting the learner with verbal and pictorial information in a coordinated way (the so-called "multimedia principle"; Mayer, 2001, 2005).

2.2. The effect of animations on learning outcome

Neither in Mayer's original theory (Mayer, 2001, 2005) nor in Schnotz's integrative model is there a basic distinction between static and animated pictures – both are examples of a pictorial presentation format. An animation can be defined as a series of rapidly changing computer screen displays suggesting movement to the viewer (Rieber & Kini, 1991). It aims at giving an exact presentation of a process or procedure to facilitate generating an adequate mental model. The supplantation framework of Salomon (1979) proposes that an animation, by dynamically displaying a process or a procedure, should be able to compensate for a student's insufficient aptitude or skill to imagine motions. Thus, the animation provides an external model for a mental representation. Mayer and Moreno (2002) showed that the basic principles of the generative theory of multimedia learning are valid for learning with animations as well as for learning with static pictures.

2.3. Cognitive load theory

When learning with dynamic or non-dynamic visualizations, the capacity of working memory sets very narrow limitations: following cognitive load theory (Chandler, 2004; Chandler & Sweller, 1991; Sweller, 1994), there are three different types of cognitive load – extraneous, intrinsic and germane cognitive load. Intrinsic cognitive load is considered as determined largely by element interactivity, i.e. the number of interacting elements in a content area, and this, therefore, cannot be manipulated by instructors and instructional designers. On the other hand, extraneous cognitive load is determined by how the information is presented – when intrinsic load is high, a high level of extraneous cognitive load can be a critical factor for successful learning (Carlson, Chandler, & Sweller, 2003). Thus, the instructional format (e.g., animations or static pictures) might influence the learning efficacy of a learning environment. By reducing extraneous cognitive load and increasing germane cognitive load – the third type of cognitive load, referring to the effort involved in the processing, construction and automation of schemas – more efficient learning may be possible.

2.4. Animations versus static pictures

There are several reasons for expecting that both visual representation formats, animations as well as static pictures, can be of benefit for learning. As for animations, one might argue that they help in mentally visualizing a process or a procedure, resulting in a reduction of cognitive load compared to a situation in which the process or the procedure has to be reconstructed from a series of static pictures. Furthermore, in static pictures often more or less abstract signaling cues like arrows or highlightings have to be interpreted and integrated with the pictorial information. This imposes even more cognitive load and can lead to misinterpretations and therefore to a deficient mental model (Lewalter, 1997). On the other hand, an animation does not provide permanent but transient information, which means that ''one views one frame at a time, and once the animation or video has advanced beyond a given frame, it is no longer available to the viewer'' (Hegarty, 2004, p. 346). This imposes – according to cognitive load theory (Chandler & Sweller, 1991) – cognitive load due to temporal limits of working memory. Furthermore, the learning efficacy of static pictures could possibly be increased by using certain key pictures that illustrate very specific moments of the process or the procedure to be learned (Catrambone & Seay, 2002; Hegarty, Kriz, & Cate, 2003). Another way to improving pictures is to use a certain level of realism (Michas & Berry, 2000).

Individual differences can influence whether static pictures or animations within a specific domain of knowledge or skills are superior: *spatial ability*, for instance, can play a critical role (Blake, 1977; Hays, 1996; Large, Beheshti, Breuleux, & Renaud, 1996; Yang, Andre, & Greenbowe, 2003) as well as prior knowledge. With higher *prior knowledge*, for instance, a learner has to invest less mental effort into learning a given topic and, thus, has more cognitive capacity left for trying to comprehend a displayed motion concerning that topic on a very detailed level (ChanLin, 2001; Höffler, 2003; Nerdel, 2003; Szabo & Poohkay, 1996).

The difference in learning from dynamic and non-dynamic pictures with retention or problem-solving tasks has been researched rather often. In particular, deeper understanding and, thus, the ability to solve advanced problems should profit from learning with animations (Mayer & Moreno, 2002). However, previous studies provide a very heterogeneous picture (e.g., Catrambone & Seay, 2002; ChanLin, 1998, 2001; Large et al., 1996; Nerdel, 2003; Wright, Milroy, & Lickorish, 1999; Yang et al., 2003).

In addition, there are several other moderators that have not been focused on in previous studies and will be surveyed in the present meta-analysis. For example, there may be a difference when the topic to be learned is explicitly depicted in the animation and when the animation is used for decorational purposes only (Rieber, 1990). Likewise, computer-based and video-based animations may differ in their advantages to their static equivalents because of different levels of realism: it can be expected that – although highly realistic pictures, like photos, are not necessarily better for learning than line drawings of the same topic (e.g., Dwyer, 1978) – highly realistic animations, like videos, can compensate or even over-compensate the disadvantage of seductive details, which are usually included in highly realistic pictures, more so than less realistic computer-based animations can do for less realistic pictures with less seductive details.

Last but not least, the specific instructional domain might also make a difference in determining the instructional effectiveness of static pictures compared to animations.

This short survey indicates that there are many theoretical arguments for the advantages of either form of visualizing processes and procedures, i.e. using static pictures and animations, and there is a wide and divergent range of research results. Thus, there are good reasons for conducting a formal meta-analysis in order to search for overall-effects and to identify moderator variables.

3. Method

In the present study, a meta-analysis was conducted following methods developed by Glass, McGaw, and Smith (1981) and Hedges and Olkin (1985). Its goals are to integrate the findings of a large number of studies, to calculate overall-effects and to identify possible moderator variables (Lipsey & Wilson, 2001). Specifically, we attempted to find overall-effects of instructional animations compared to static pictures on learning outcomes. Furthermore, factors or variables moderating the effect size are to be identified. A meta-analysis is traditionally conducted in three main steps: (1) location and selection of appropriate studies, (2) coding of study features and calculating effect sizes, and (3) statistically analyzing effect sizes and the influence of study features.

3.1. Location and selection of appropriate studies

To identify studies comparing the effects of animations versus static pictures, the computerized databases *SCI* and *SSCI* (1993–2004), *ERIC* (1966–2004), *PsycInfo* (1887–2004) and *Psyndex* (1977–2004) were searched. A range of combinations of descriptors such as "animation", "dynamic picture", "dynamic image", "still picture", "still image", "static picture", "static image", "motion", "steps", "simulation", etc. were used. The large number of hits was screened on the basis of their abstracts, if available. When in doubt of whether to exclude an article, the full document was retrieved. Apart from the articles found in databases, cross-references from identified articles helped to find additional studies. Especially useful were the reviews of Alesandrini (1982), Bétrancourt and Tversky (2000), Large (1996), Milheim (1993), Park and Hopkins (1993), and Tversky et al. (2002). To counteract a possible publication bias, we tried to include some unpublished dissertations, diploma theses, and conference proceedings as well. These were, however, difficult to detect and often did not fulfill the methodological criteria listed below.

The search located 57 articles for closer examination. In the next step, these studies were checked whether they fulfilled the following four criteria for being included in the meta-analysis: the study (1) compared animated with static displays, (2) did not mix both types of visualization (otherwise the versions would not be comparable), (3) had no (or only minimal) interactivity within the animation (e.g., options of changing parameters so that the animation would not be comparable to static pictures), and (4) investigated static pictures and animations that are roughly equivalent concerning the specific content presented to the learner. Based on these criteria, 25 of 57 articles were excluded from the meta-analysis. Five other studies had to be excluded because they did not specify the basic statistics needed for computing effect sizes or did not meet minimal statistical standards. However, this criterion was handled fairly liberally, as it is customary in meta-analyses (Bangert-Drowns, 1986). One article had to be excluded because it presented the same statistical material that had previously been published in another article (ChanLin, 2000). In the end, 26 studies were included in the present meta-analysis.

3.2. Coding of study features

While all included studies examined the differences in learning outcomes between static pictures and animations, this was seldom their only focus. Each study, of course, had different emphases. These circumstances made it necessary to code a multitude of study features to be able to identify as many moderator variables as possible that might account for variation in effect sizes. As a result, the following features were coded:

- Three coded variables focused on *features of the animation version*:
  (1) Distinction between video-based animations and computer-based animations.
  (2) Differentiation between four rated levels of realism of the animation (schematic, rather simple, rather realistic, photo-realistic [= video]).
  (3) Distinction between representational animations and decorational animations. Adapted from the denotation of Carney and Levin (2002), in a representational animation the topic to be learned is explicitly depicted in the animation, whereas in a decorational animation the primary instructional function is to motivate the learner.
- Two variables focused on *additional features*:
  (4) It could make a difference in learning efficacy if visualizations are annotated by coherent text. It should be mentioned that in all analyzed studies care was taken that neither version of visualization included more information, i.e. whenever static pictures were accompanied by annotating text, the same was true for the animation.
  (5) Sometimes visualizations were provided with signaling cues like arrows and highlighting.
- Two variables coded the *characteristics of the learning task*:
  (6) One variable classified the instructional domain, e.g., biology, mathematics, military, etc.
  (7) The type of knowledge (as the specific goal of learning) was classified into three categories: procedural-motor knowledge, declarative knowledge and problem-solving knowledge.
- One variable followed a remark of Tversky et al. (2002) that often both versions, i.e. animation and static pictures, differed in *learners' time on task* and therefore were not comparable:
  (8) The amount of time the learners worked with either version was coded when specified.
- Finally, substantial study features were listed:
  (9) Year of publication.
  (10) Sample sizes.
  (11) Sample characteristics (students, undergraduates, recruits, adults, etc.).

3.3. Calculation and analysis of effect sizes

The effect sizes were calculated as Cohen's d in the modified form of Hedges and Olkin (1985): the mean of one group was subtracted from the mean of the other group, divided by the pooled standard deviation. When means or standard deviations were not reported, formulas for calculating d from t or F statistics (Glass et al., 1981) were used, or an estimation of d from χ^2 (Cohen, 1966) was obtained. As the effect of animations versus static pictures was not the main target of most studies, the relevant statistics were often not listed separately. Instead, mean differences between those test conditions which were the focus of the specific study were provided. Thus, the method of Glass et al. (1981) was chosen to integrate all given effect sizes in the meta-analysis and to handle each pair-wise comparison as an independent study. As a result, 26 studies provided 76 effect sizes, a positive d indicating an advantage of the animation over the static pictures, a negative d indicating an advantage of the static pictures over the animation. We followed Lipsey and Wilsons's (2001) suggestion to identify outliers which are located more than 3 standard deviations from the mean effect size and recoded them to the value of mean effect size ± 3 standard deviations (it turned out that this had to be done in two pair-wise comparisons derived from a study of Blake, 1977).

A correction for small sample bias in effect-size estimation (Hedges & Olkin, 1985) was calculated. Furthermore, to avoid over-representing studies with many pair-wise comparisons, we calculated weighted effect sizes. This commonly used strategy in meta-analysis gives greater weight to pair-wise comparisons with larger samples, assuming that larger samples yield better estimates of population parameters (Bangert-Drowns, Hurley, & Wilkinson, 2004). Thus no study received a disproportionally large weight because of a large number of effect sizes derived from that study. In fact, each pair-wise comparison is based on fewer subjects than the whole study and accordingly achieves less weight. Thus, the sum of weights of all pair-wise comparisons derived from one study does not exceed the total weight the study should have based on its sample size.[1] To reach this goal, each effect size was weighted by the inverse of the effect size's standard error (Hedges & Olkin, 1985).

We adopted a random-effects model for calculating estimates of mean effect sizes and 95% confidence intervals around these estimates (using the software Zumastat; Jaccard, 2006). In comparison to fixed-effects models, random-effects models tend to yield confidence intervals closer to their nominal width (e.g., Quintana & Minami, 2006) and are therefore recommended by a number of authors (e.g., Erez, Bloom, & Wells, 1996; Hunter & Schmidt, 2000). The random-effects model was applied for calculating both overall analyses and moderator analyses. In case of moderating variables consisting of more than two categories, Bonferroni-corrected pair-wise comparisons were conducted (setting alpha at 0.05 and using the Holm-modified Bonferroni procedure[2]; Holm, 1979; for an example of this approach, see Ginns, 2005).

4. Results

4.1. Characteristics of the sample

Twenty-six studies yielding 76 pair-wise comparisons of learning outcome differences between instructional animations and static pictures were included in the meta-analysis. The selected studies were published between 1973 and 2003, but only three of them were published before 1980, when the technological potentials of computer-based

[1] Some studies, which are indicated in Table 1, used multiple learning outcome measures and/or multiple comparisons of different experimental groups with the same control group. Adapting a suggestion of Hedges and Olkin (1985, p. 206), the weights of effect sizes derived from these studies were additionally adjusted according to the number of those non-independent pair-wise comparisons. For example, Craig, Gholson, and Driscoll (2002) used *four* different multiple learning outcome measures and compared *two* different static picture versions to the same animated version. Therefore, eight different pair-wise comparisons were included in the meta-analysis. As these comparisons consisted of the same 30 subjects, their sample size had to be adjusted to $30/8 = 3.75$ each, thereby assuring that no comparison gained inappropriate weight. Another way to handle dependency problems of this type would have been to apply a hierarchical method of excluding specific pair-wise comparisons from the meta-analysis (e.g., Ginns, 2005). We chose to adjust sample sizes and, thus, to adjust the weights of pair-wise comparisons instead of the exclusion procedure in order not to exclude too much information that might be informative for moderator analyses from our analysis.

[2] Holm's (1979) method compares the most significant pair-wise comparison to a critical value of alpha divided by k, the number of comparisons. Afterwards, the second most significant comparison is compared to a critical value of alpha/($k - 1$), and so on, until the first non-rejected null-hypothesis occurs.

learning environments were very restricted. Only 6 out of 26 studies used video-clips in contrast to static pictures; the other studies compared computer-based animations with static pictures.

The sample sizes for calculating mean learning outcomes ranged from 21 to 263; the mean sample size was 55 and the median 40. In 17 studies (65%) the participants were undergraduates and in 7 studies (27%) high school students. In one study the participants were adults and in one study recruits. Table 1 lists all included studies with selected study features.

4.2. Overall-effects of animations

If there were no differences in learning outcomes between animations and static pictures one could assume that the effect sizes would show an approximate normal distribution with an expected mean of zero. However, the meta-analysis resulted in 21 pair-wise comparisons with a statistically significant advantage of the animation, in only 2 pair-wise comparisons the static pictures are significantly superior. A possible publication bias appears to be unlikely, as there are 53 pair-wise comparisons in which the difference between animation and static pictures is not statistically significant. The *fail-safe N* (Rosenthal, 1979; modified by Orwin, 1983) – that is the fictitious number of non-significant comparisons that would be necessary to be integrated in the meta-analysis in order to cause the overall-effect to be no longer existent – is calculated to be 2736. Following Rosenthal (1991), a publication bias seems reasonably implausible when the fail-safe N exceeds the quintuple of the total number of included effect sizes plus 10, which is clearly the case (here: $5 \times 76 + 10 = 390$).

Focusing on the distribution of derived effect sizes, 54 out of 76 pair-wise comparisons (71%) are positive indicating an advantage of animations over static pictures. The mean weighted effect size (corrected according to Hedges & Olkin, 1985) is $d = 0.37$ standard deviations with a 95% confidence interval of 0.25–0.49.

Fig. 1 shows the histogram of weighted effect sizes. It suggests that the effect-size distribution does not represent a single and homogeneous population of effect sizes but rather reflects differences in study features.[3] An overall homogeneity test (Hedges & Olkin, 1985) confirms that the set of effect sizes is heterogeneous: $Q_{\text{total}} = 425.09$, df = 75, $p < 0.001$. Hence, a detailed analysis of moderator variables is warranted.

4.3. Impact of moderator variables

Following the overall analysis of effect sizes, the impact of potential moderator variables was investigated. In these analyses, the following variables were included: (1) role of animation (representational, decorational), (2) type of requested knowledge (procedural-motor, declarative, problem-solving knowledge), (3) type of animation (video-based, computer-based), (4) level of realism (schematic, rather simple, rather realistic, photo-realistic), (5) annotating text (included, not included), (6) cues in static pictures (included, not included), and (7) instructional domain (biology, mathematics, military, etc.). Learners' time on task, unfortunately, could not be included, as there were only seven studies indicating this variable. Table 2 displays number of effect sizes, weighted mean effect sizes and confidence intervals within the categories of the included variables.

The analyses showed significant moderating effects which will be presented and discussed in the following sections.

4.3.1. Role of animation

It is reasonable to expect that it makes a difference whether the topic to be learned is explicitly depicted in the animation or not, i.e. whether the animation has a representational rather than a decorational function (Carney & Levin, 2002). The results indicate (Table 2) that representational animations are significantly superior to representational static pictures ($d = 0.40$; 95% confidence interval, 95% CI, 0.26–0.53) whereas decorational animations are not significantly superior to decorational static pictures ($d = -0.05$, 95% CI −0.37 to 0.27). The difference between the two mean weighted effect sizes reaches statistical significance ($z_{\text{contrast}} = 3.86$, $p < 0.001$).

[3] There appears to be an outlier at $d = 2.63$. Note that a correction for outliers has already been performed at the raw effect-size level. Therefore, it is not reasonable to conduct a second adjustment on the weighted effect-size level. In fact, such a second adjustment to the value of mean weighted effect size +3 standard deviations (according to Lipsey & Wilson, 2001) would not affect the results.

Table 1
Selected features of 26 primary studies yielding 76 pair-wise comparisons of animations versus static pictures

Study	Total sample size N	Sample size n of pair-wise comparison	Adjusted sample size of pair-wise comparison	Weighted effect size d	Sample	Instructional domain	Type of animation	Level of realism	Cues	Text	Role of animation	Type of requested knowledge
Baek & Layne, 1988	119[a]	46	46.0	0.58	Students	Mathematical rule for average speed	Computer-based	2	n/a	Yes	Representational	Problem-solving
		38	38.0	0.35	Students		Computer-based	2	n/a	Yes	Representational	Problem-solving
Blake, 1977	84[b]	28	28.0	1.25[r]	Undergraduates	Movement patterns of chessmen	Video-based	4	Yes	No	Representational	Declarative
		28	28.0	0.10	Undergraduates		Video-based	4	Yes	No	Representational	Declarative
		28	28.0	1.25[r]	Undergraduates		Video-based	4	No	No	Representational	Declarative
		28	28.0	0.97	Undergraduates		Video-based	4	No	No	Representational	Declarative
Catrambone & Seay, 2002, Exp. 2	188[c,d]	188	94.0	0.74	Undergraduates	Computer algorithms	Computer-based	1	No	Yes	Decorational	Problem-solving
		188	94.0	−1.19	Undergraduates		Computer-based	1	No	Yes	Decorational	Problem-solving
ChanLin, 1998	135[a,c,e]	50	25.0	0.37	Undergraduates	Recombinant DNA technology	Computer-based	3	Yes	Yes	Representational	Declarative
		50	25.0	−0.38	Undergraduates		Computer-based	3	Yes	Yes	Representational	Problem-solving
		40	20.0	0.00	Undergraduates		Computer-based	3	Yes	Yes	Representational	Declarative
		40	20.0	0.11	Undergraduates		Computer-based	3	Yes	Yes	Representational	Problem-solving
ChanLin, 2001	357[a,c,e]	142	71.0	1.13	Students	Forces in physics	Computer-based	2	Yes	Yes	Decorational	Declarative
		142	71.0	0.97	Students		Computer-based	2	Yes	Yes	Decorational	Problem-solving
		92	46.0	−0.86	Students		Computer-based	2	Yes	Yes	Decorational	Declarative
		92	46.0	−1.13	Students		Computer-based	2	Yes	Yes	Decorational	Problem-solving
Craig, Gholson, & Driscoll, 2002, Exp. 1	135[b,c,f]	30	3.75	0.01	Undergraduates	Lightning formation	Computer-based	3	Yes	Yes	Representational	Declarative
		30	3.75	0.06	Undergraduates		Computer-based	3	No	Yes	Representational	Declarative
		30	3.75	−0.01	Undergraduates		Computer-based	3	Yes	Yes	Representational	Declarative
		30	3.75	0.04	Undergraduates		Computer-based	3	No	Yes	Representational	Declarative
		30	3.75	0.02	Undergraduates		Computer-based	3	Yes	Yes	Representational	Declarative
		30	3.75	0.08	Undergraduates		Computer-based	3	No	Yes	Representational	Declarative
		30	3.75	−0.01	Undergraduates		Computer-based	3	Yes	Yes	Representational	Problem-solving
		30	3.75	0.03	Undergraduates		Computer-based	3	No	Yes	Representational	Problem-solving
Hays, 1996	116[a,c,g]	77	77.0	0.11	Students	Diffusion	Computer-based	3	n/a	Yes	Representational	Declarative
		67	67.0	1. 05	Students		Computer-based	3	n/a	Yes	Representational	Problem-solving
Höffler, 2003	115[c,f,h]	59	29.5	−0.19	Undergraduates	Photosynthesis	Computer-based	3	Yes	Yes	Representational	Declarative
		59	29.5	−0.15	Undergraduates		Computer-based	3	Yes	Yes	Representational	Problem-solving
Kaiser, Proffitt, & Anderson, 1985, Exp. 1	105[i]	51	51.0	1.72	Undergraduates	Natural and anomalous trajectories	Video-based	4	Yes	No	Representational	Problem-solving

Lai, 2000	316[c,j,k]	126	63.0	−1.12	Undergraduates	Programming concepts of *Quick BASIC*	Computer-based	2	Yes	Yes	Decorational	Declarative
		126	63.0	−0.04	Undergraduates		Computer-based	2	Yes	Yes	Decorational	Declarative
Lewalter, 2003	60[a,c,f,l]	40	20.0	0.00	Undergraduates	Gravitation	Computer-based	3	Yes	Yes	Decorational	Declarative
		40	20.0	0.18	Undergraduates		Computer-based	3	Yes	Yes	Decorational	Problem-solving
McCloskey & Kohl, 1983, Exp. 2	72	72	72.0	−0.53	Undergraduates	Natural and anomalous trajectories	Video-based	4	Yes	No	Representational	Problem-solving
Michas & Berry, 2000, Exp. 1	75[b,j]	30	15.0	0.61	Undergraduates	How to bandage a hand	Video-based	4	No	No	Representational	Procedural-motor
		30	15.0	0.52	Undergraduates		Video-based	4	Yes	No	Representational	Procedural-motor
Nerdel, 2003, Exps. 2 and 3	131[c,f,h]	55	27.5	0.30	Students	Respiratory chain	Computer-based	3	Yes	Yes	Representational	Declarative
		55	27.5	0.38	Students		Computer-based	3	Yes	Yes	Representational	Problem-solving
	150[c,f,h]	60	30.0	0.04	Students	Photosynthesis	Computer-based	3	Yes	Yes	Representational	Declarative
		60	30.0	−0.20	Students		Computer-based	3	Yes	Yes	Representational	Problem-solving
Nicholls & Merkel, 1996, Exp. 1	44[m]	21	21.0	0.27	Undergraduates	Nitrogen cycle	Computer-based	n/a	n/a	Yes	Representational	Problem-solving
Rieber, 1989	192[a,c,n]	32	8.0	−0.02	Students	Newton's laws of motion	Computer-based	1	Yes	No	Representational	Problem-solving
		32	8.0	0.06	Students		Computer-based	1	Yes	No	Representational	Problem-solving
		32	8.0	0.05	Students		Computer-based	1	Yes	No	Representational	Declarative
		32	8.0	0.01	Students		Computer-based	1	Yes	No	Representational	Declarative
		32	8.0	0.14	Students		Computer-based	1	Yes	No	Representational	Problem-solving
		32	8.0	0.02	Students		Computer-based	1	Yes	No	Representational	Problem-solving
		32	8.0	0.07	Students		Computer-based	1	Yes	No	Representational	Declarative
		32	8.0	0.04	Students		Computer-based	1	Yes	No	Representational	Declarative
		32	8.0	0.08	Students		Computer-based	1	Yes	Yes	Representational	Problem-solving
		32	8.0	0.11	Students		Computer-based	1	Yes	Yes	Representational	Problem-solving
		32	8.0	0.04	Students		Computer-based	1	Yes	Yes	Representational	Declarative
		32	8.0	0.07	Students		Computer-based	1	Yes	Yes	Representational	Declarative
		32	8.0	−0.01	Students		Computer-based	1	Yes	Yes	Representational	Problem-solving
		32	8.0	−0.11	Students		Computer-based	1	Yes	Yes	Representational	Problem-solving
		32	8.0	−0.09	Students		Computer-based	1	Yes	Yes	Representational	Declarative
		32	8.0	−0.01	Students		Computer-based	1	Yes	Yes	Representational	Declarative
Rieber, 1990	119[a]	79	79.0	1.32	Students	Newton's laws of motion	Computer-based	1	Yes	Yes	Representational	Problem-solving
Rieber, 1991	70[c,f]	70	35.0	0.73	Students	Newton's laws of motion	Computer-based	1	Yes	Yes	Representational	Declarative
		70	35.0	1.48	Students		Computer-based	1	Yes	Yes	Representational	Problem-solving
Rieber, Boyce, & Assad, 1990	141[a,j]	31	31.0	0.04	Undergraduates	Newton's laws of motion	Computer-based	1	Yes	Yes	Representational	Problem-solving
		31	31.0	−0.21	Undergraduates		Computer-based	1	Yes	Yes	Representational	Problem-solving
		31	31.0	0.22	Undergraduates		Computer-based	1	Yes	Yes	Representational	Problem-solving

(*continued on next page*)

Table 1 (*continued*)

Study	Total sample size *N*	Sample size *n* of pair-wise comparison	Adjusted sample size of pair-wise comparison	Weighted effect size *d*	Sample	Instructional domain	Type of animation	Level of realism	Cues	Text	Role of animation	Type of requested knowledge
Rigney & Lutz, 1976	40[c,o]	40	10.0	0.21	Undergraduates	Concepts of electrochemistry	Computer-based	n/a	n/a	Yes	Decorational	Declarative
		40	10.0	0.15	Undergraduates		Computer-based	n/a	n/a	Yes	Decorational	Declarative
		40	10.0	0.21	Undergraduates		Computer-based	n/a	n/a	Yes	Decorational	Declarative
		40	10.0	0.18	Undergraduates		Computer-based	n/a	n/a	Yes	Decorational	Declarative
Spangenberg, 1973, Exps. 1 and 2	120	40	40.0	1.44	Recruits	Disassembly of a machine gun	Video-based	4	No	No	Representational	Procedural-motor
		40	40.0	1.10	Recruits		Video-based	4	Yes	No	Representational	Procedural-motor
		40	40.0	1.10	Recruits		Video-based	4	No	No	Representational	Procedural-motor
Spotts & Dwyer, 1996	63[j]	41	41.0	0.76	Undergraduates	Blood flow in human heart	Computer-based	2	No	Yes	Representational	Declarative
Swezey, 1991	120[p]	120	120.0	0.00	Undergraduates	Functions of a diesel engine	Video-based	4	No	No	Decorational	Declarative
Szabo & Poohkay, 1996	174[a]	117	117.0	2.63	Undergraduates	Triangle-construction using a compass	Computer-based	2	Yes	Yes	Representational	Declarative
Wright et al., 1999, Exp. 1	60[a,c,q]	40	20.0	0.24	Adults	British history	Computer-based	n/a	n/a	Yes	Decorational	Declarative
		40	20.0	−0.09	Adults		Computer-based	n/a	n/a	Yes	Decorational	Declarative
Yang et al., 2003	263	263	263.0	1.76	Undergraduates	Electro-chemical principles in a flashlight	Computer-based	2	Yes	Yes	Representational	Declarative

[a] Another test condition which was not considered in the meta-analysis was text only.
[b] Two different static versions (with/without cues) were compared to one animated version. *n* was adjusted accordingly.
[c] Sample sizes *n* were adjusted due to the inclusion of multiple dependent variables.
[d] "Near transfer" and "far transfer" tasks were given to the same subjects. Results were averaged.
[e] Experienced and novice learners were distinguished and asked to answer declarative and problem-solving tasks.
[f] Declarative and problem-solving tasks were given to the same subjects and were analyzed separately.
[g] Problem-solving knowledge was assessed 1 week after testing as "long-term comprehension".
[h] Other test conditions were interactive animations (simulations), which were not regarded in the meta-analysis.
[i] The authors do not state why they used only 51 of their 105 participants to analyze differences between "motion and no-motion condition".
[j] Additional test conditions were not regarded in the meta-analysis.
[k] The same test was used as posttest and retention test 1 week later.
[l] Two different problem-solving tests were averaged.
[m] The authors state that "due to a miscommunication" another group of subjects should not be considered for further analyses.
[n] Declarative and problem-solving tasks were further divided in "near" and "far". All conditions were analyzed separately.
[o] Four different declarative tests were given to the same subjects and were analyzed separately.
[p] Different knowledge tests were used, but only one ("conceptual knowledge") was clearly relatable to the categories of the meta-analysis and, thus, further analyzable.
[q] Two different knowledge tests existed and were analyzed separately.
[r] The raw effect size has been adjusted beforehand because it was larger than the mean raw effect size +3 standard deviations.

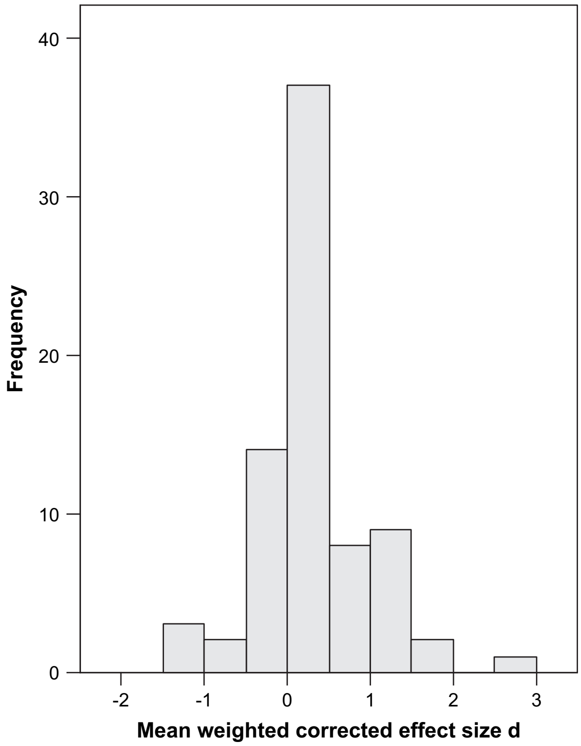

Fig. 1. Distribution of weighted effect sizes.

4.3.2. Type of requested knowledge

In different studies – as well as sometimes within the same study – different types of knowledge were tested: in most cases ($N = 40$ comparisons), declarative knowledge had to be learned. Often ($N = 31$) the pair-wise comparisons focused on deeper comprehension allowing learners to solve problems (problem-solving knowledge). A third type of requested knowledge is procedural-motor knowledge. This type of knowledge, tested in five comparisons, for example requested the trained capability to reconstruct a machine gun (Spangenberg, 1973). Procedural-motor knowledge (Table 2) has the largest mean weighted effect size in favor of animations ($d = 1.06$, 95% CI 0.72–1.40). When declarative knowledge was requested, the mean effect size is $d = 0.44$ (95% CI 0.27–0.57), whereas for problem-solving knowledge the mean effect size is $d = 0.24$ (95% CI 0.04–0.44). Contrasts between the different types of requested knowledge indicate two significant differences: the mean effect size involving procedural-motor knowledge exceeds the mean effect size involving problem-solving knowledge ($z_{contrast} = 4.16$, $p < 0.001$) as well as the mean effect size involving declarative knowledge ($z_{contrast} = 3.41$, $p = 0.001$). Note that these results may be influenced by some confounding of type of knowledge with role of animation (see above), as all five visualizations requesting procedural-motor knowledge were also representational, whereas a considerable number of visualizations requesting declarative knowledge or problem-solving knowledge were decorational (12 out of 40, 5 out of 31, respectively). Mean effect sizes involving declarative or problem-solving knowledge do not differ significantly ($z_{contrast} = 1.41$, $p = 0.159$).

4.3.3. Type of animation

In 12 comparisons of animations and static pictures, the animations were video-based. For these animations (Table 2), the mean weighted effect size is $d = 0.76$ (95% CI 0.39–1.13), whereas for computer-based animations the mean effect size is only $d = 0.36$ (95% CI 0.25–0.46), the difference being statistically significant ($z_{contrast} = 2.06$, $p = 0.039$). Again, this result may be influenced by some confounding of type of animation with role of animation (see above): nearly all, 11 out of 12, video-based animations (92%), whereas only 48 out of 64 computer-based animations (75%) were used in a representational role. When looking at the representational animations separately, the effect-size difference between video-based animations ($d = 0.85$, 95% CI 0.47–1.23) and computer-based animations ($d = 0.68$,

Table 2
Mean weighted effect sizes and confidence intervals for moderator categories

Moderator variable	Number of effect sizes	Mean weighted d	95% CI for d
Role of animation			
Decorational animation	17	−0.05	−0.37 to 0.27
Representational animation	59	0.40	0.26–0.53
Type of requested knowledge			
Procedural-motor	5	1.06	0.72–1.40
Declarative	40	0.44	0.27–0.57
Problem-solving	31	0.24	0.04–0.44
Type of animation			
Computer-based	64	0.36	0.25–0.46
Video-based	12	0.76	0.39–1.13
Level of realism of animation			
Schematic	24	0.24	0.07–0.42
Rather simple	11	0.47	−0.24 to 1.17
Rather realistic	22	0.17	−0.01 to 0.35
Photo-realistic	12	0.76	0.39–1.13
Annotating text			
Included	59	0.35	0.21–0.49
Not included	17	0.39	0.16–0.62
Signaling cues in static pictures			
Included	52	0.33	0.16–0.49
Not included	13	0.47	0.18–0.76
Instructional domain			
Biology	12	0.13	−0.09 to 0.34
Physics	39	0.28	−0.15 to 0.41
Chemistry	7	0.75	−0.33 to 1.16
Mathematics	5	0.62	−0.57 to 1.81
Military	3	1.21	0.82–1.60
Other	10	0.32	−0.12 to 0.76

95% CI 0.57–0.79) vanishes ($z_{contrast} = 0.85$, $p = 0.395$). On the other hand, 16 out of 64 computer-based animations (25%) are decorational with $d = -0.05$ (95% CI −0.39 to 0.29), and the effect size of the one video-based animation, which is – as it turns out – also decorational (Swezey, 1991), is zero. Thus, the *overall* smaller effect size of computer-based animations compared to video-based animations reflects, at least partly, that nearly all video-based animations are at the same time representational whereas a quarter of the computer-based animations are at the same time decorational: when these decorational animations with effect sizes at zero are excluded, the difference between video-based and computer-based animations is no longer statistically significant.

4.3.4. Level of realism

Video-based animations are by definition photo-realistic (level 4). As computer-based animations can have different grades of realism, we coded them on three levels: schematic (level 1), rather simple (level 2), and rather realistic (level 3). Between the four levels of realism (level 1: $d = 0.24$, 95% CI 0.07–0.42; level 2: $d = 0.47$, 95% CI −0.24 to 1.17; level 3: $d = 0.17$, 95% CI −0.01 to 0.35; level 4: $d = 0.76$, 95% CI 0.39–1.13; Table 2), there are two statistically significant effect-size differences, which are in line with the difference between video-based and computer-based animations as shown in Section 4.3.3: the contrast, adjusted according to Holm–Bonferroni, of level 3 (rather realistic) and level 4 (photo-realistic) is significant ($z_{contrast} = 2.81$, $p = 0.005$) as well as the contrast of level 1 (schematic) and level 4 (photo-realistic; $z_{contrast} = 2.47$, $p = 0.014$). The results are in line with the difference between video-based and computer-based animations as shown in Section 4.3.3.

4.3.5. Annotating text

In 59 comparisons of animations and static pictures, the visualizations had annotating text or narration, in 17 comparisons they had not. In both cases (Table 2), animations are superior to static pictures (with text: $d = 0.35$, 95% CI 0.21–0.49; without text: $d = 0.39$, 95% CI 0.16–0.62). However, the effect-size difference is not statistically significant in the expected direction ($z_{\mathrm{contrast}} = 0.29$, $p = 0.386$).

4.3.6. Signaling cues in static pictures

In 52 comparisons of animations and static pictures, the pictures had signaling cues like arrows and highlighting, in 13 comparisons the pictures did not have such signaling cues. Animations turn out to be superior to static pictures (Table 2) when the pictures do not have signaling cues ($d = 0.47$, 95% CI 0.18–0.76) – and seemingly less so when static pictures do have signaling cues ($d = 0.33$, 95% CI 0.16–0.49). However, the effect-size difference is not statistically significant in the expected direction ($z_{\mathrm{contrast}} = 0.82$, $p = 0.206$).

4.3.7. Instructional domain

This moderator variable is characterized by a rather rough categorization of the instructional domains based on the particular learning environment from which the effect sizes were calculated. Clear-cut encodings were sometimes difficult to make, which might be a reason for obtaining rather large confidence intervals in some cases (Table 2). Contrasts, adjusted according to Holm–Bonferroni, show that visualizations in the domain of military, investigated in three cases only (Spangenberg, 1973; $d = 1.21$, 95% CI 0.82–1.60), have a statistically significant larger mean weighted effect size than those in the domains of physics ($d = 0.28$, 95% CI −0.15 to 0.41; $z_{\mathrm{contrast}} = 4.39$, $p < 0.001$), biology ($d = 0.13$, 95% CI −0.09 to 0.34; $z_{\mathrm{contrast}} = 4.73$, $p < 0.001$), and "other" domains ($d = 0.32$, 95% CI −0.12 to 0.76; $z_{\mathrm{contrast}} = 2.92$, $p = 0.004$). However, the three visualizations from the military domain are also classified to be representational (role of animation), to focus on procedural-motor knowledge (type of requested knowledge), and to be video-based (type of animation). Thus, the confounding of "military domain" with these features (see above) might, at least to a certain amount, account for the stronger effect size.

5. Summary and discussion

In the present meta-analysis, 76 pair-wise comparisons out of 26 studies comparing the instructional effectiveness of animations with static pictures were included. Five comparative questions will be used to summarize the results.

5.1. Are animations better than static pictures in general?

Ignoring moderator variables, a clear advantage of non-interactive animations compared to static pictures was observed in the present meta-analysis: the mean weighted effect size of $d = 0.37$ (95% CI 0.25–0.49) indicates a small to medium effect (Cohen, 1988). Tallmadge (1977) even considers effect sizes of $d = 0.25$ to $d = 0.33$ as "educationally significant" (p. 34). Taking into account that weighted effect sizes were calculated and that a publication bias seems to be unlikely (based on fail-safe N calculations according to Rosenthal, 1979), other than purely statistical reasons like sampling error may account for the present findings. Given that this assumption is valid, the results of the present meta-analysis seem to contradict the mainstream of contemporary research on instructional animations according to which non-interactive animations are usually not regarded as universally helpful for learning (Bétrancourt & Tversky, 2000). To the contrary, it is reasonable to assume that animations are able to take advantage of their specific characteristics under specific circumstances, which can then result in rather large effect sizes. Although many primary studies did not find a significant advantage of animations over static pictures, an overall mean weighted effect size of $d = 0.37$ indicates that there seem to be instructional situations in which the particular benefits of animations arise. In unison with other authors (e.g., Mayer & Moreno, 2002; Milheim, 1993; Tversky et al., 2002; Weiss, Knowlton, & Morrison, 2002) the moderator analyses discussed below are an attempt to identify some of them.

5.2. Are representational animations better than decorational animations?

One clearly identified moderating variable in the present meta-analysis is the instructional role of animation. It makes a difference whether the topic to be learned is explicitly depicted in the animation or not, or, in other words, whether the animation has a representational rather than a decorational function (Carney & Levin, 2002): representational animations are far more superior to static pictures than are decorational animations (mean weighted effect size of $d = 0.40$ versus $d = -0.05$). Hence, animations seem to be especially useful when the motion depicted in the animation is the content to be learned. However, across the primary studies of the present meta-analysis, the classification variable ‘‘role of animation’’ co-varies with other classification variables so that the significance of representational-animation effect sizes and the non-significance of decorational-animation effect sizes, as appropriate, have to be taken into account when interpreting the moderator effects of other classification variables.

Examples for a successful use of a representational animation are the study of Michas and Berry (2000), in which the exact procedure of how to bandage a hand is taught, and the study of Yang et al. (2003), in which students had to learn about the motions of electrons inside a flashlight battery. One interpretation of this finding is that an animation, when representational, facilitates generating a mental model of the motion to be learned by providing a prototype: ‘‘An animation is likely to be useful when the learning material entails motion, trajectory or change over time so that the animation helps to build a mental model of the dynamics’’ (Bétrancourt & Tversky, 2000).

Whenever an animation is only used for decorational purposes, a learner’s mental model ‘‘in motion’’ does not seem to be necessary for understanding. In the study of Lai (2000), for example, in which concepts of programming with *Quick BASIC* were taught via an animated comic figure, the animation did not have an advantage over static pictures. To the contrary, in such a case the animation may distract the learner’s attention from the actual topic to be learned: thus, following Cognitive Load Theory (Chandler & Sweller, 1991), the animation would impose *extraneous cognitive load* on the learner, which burdens the capacity of working memory unnecessarily. In other words, the decorational animation may offer *seductive details* (Harp & Mayer, 1998). Although decorational visualizations are often assumed to have specific effects on a learner’s motivation (Levin et al., 1987), this effect seems to play only a minor role in the present context.

5.3. Are animations better for acquiring procedural-motor knowledge rather than declarative knowledge or problem-solving knowledge?

The results of the present meta-analysis reveal greater benefits of animations when procedural-motor knowledge rather than problem-solving knowledge or declarative knowledge is requested. The effect-size difference between problem-solving knowledge and declarative knowledge, however, falls short of being significant (a future meta-analysis, including incoming primary studies, should repeat the test).

The effect-size difference between procedural-motor knowledge and problem-solving knowledge is at least partly due to the fact that all effect sizes concerning procedural-motor knowledge are exclusively based on representational animations and pictures whereas effect sizes concerning problem-solving knowledge are based on representational as well as decorational animations and pictures. These results are partly surprising, as animations are commonly viewed as especially beneficial for the comprehension of processes or procedures – whereas for learning of simple facts static pictures should suffice. Weiss et al. (2002) recommend that, for teaching of procedures, an ‘‘animation might be useful in helping your audience understand the steps in the procedure’’ (p. 474). Accordingly, the superiority of animations when procedural-motor knowledge is requested is in line with this recommendation. On the other hand, a similar superiority would have been expected for animations when requesting problem-solving knowledge (e.g., Mayer & Moreno, 2002). The present results allow one to question this common prescription – in many cases, static pictures may suffice to learn not only simple facts, but even to gain deeper understanding.

5.4. Are computer-based animations better than video-based animations?

Concerning the question whether animations should be computer-based rather than video-based – which leads directly to the question of an adequate level of realism – the pattern of results of the present meta-analysis is rather inconclusive. Although, at first sight, video-based animations seem to be superior to computer-based animations, the advantage of highly realistic (video-based) animations can, at least to a certain amount, be attributed to a confounding with the role of animation: all video-based animations are also representational whereas a substantial portion of computer-based animations is decorational.

Thus, compared to learning with static pictures, animations with lower levels of realism do not necessarily result in smaller effect sizes, which is consistent with Tversky et al. (2002): ‘‘Animations should lean toward the schematic and away from the realistic’’ (p. 258). In so doing, learners should be able to concentrate on essential contents of the animation – ‘‘sufficiently complex to convey the important information within it, yet simple enough to be easily understood’’ (Milheim, 1993, p. 173; see also Lowe, 1999, 2003). Rieber (1994) demanded as well that animations should directly indicate the crucial aspects of the topic. In his studies (Rieber, 1990, 1991), he used a chunking-strategy: "By partitioning the animations to discrete steps the new information was presented step-by-step". In addition, Rieber used animations with a very low level of realism. Rieber (1994) remarked that his chunking-strategy is only one of several possibilities of cueing in animations (p. 176). Following Tversky et al. (2002), such a cueing strategy could include narrations, arrows or graphic accentuations. Finally, the critical point seems to be to reduce learners' cognitive load by excluding as many elements as possible, which disturb attention and – thus – increase the amount of extraneous cognitive load (Sweller, 1994).

5.5. Can static pictures be improved?

The present meta-analysis suggests that instructional animations are, in general, superior to static pictures with respect to learning outcome. Based on this result, an important question arises: can static pictures be designed to compensate somewhat for their disadvantages? Interestingly, we did not find that the presence or absence of annotating text in animations and pictures accounts for differences in effect size: the superiority of animations to static pictures does not vary significantly. While Mayer's "multimedia principle" (Mayer & Gallini, 1990) states that static pictures and text are better for learning than text alone, one could have assumed that static pictures and text should have been better than pictures alone. This is, however, not the case for the primary studies included in the present meta-analysis.

Likewise, the effect size in favor of animations does not vary significantly depending on the presence or absence of signaling cues such as arrows and highlighting in static pictures. Concerning static pictures one might have expected that there is no necessity to point out their most important elements, as the viewer normally has sufficient time to find them by him- or herself. On the other hand, concerning animations, especially non-interactive ones, cues might have been expected to be reasonable and "conducive" (as an antonym to "seductive" details; Harp & Mayer, 1998): in principle, there is the risk of missing critical aspects because of the transient character of the animation with high demands on working-memory capacity (Ainsworth & VanLabeke, 2004). However, the present meta-analysis does not provide evidence for such a hypothesis. Future meta-analytic research based on a larger number of incoming primary studies is needed for clarifying this question.

5.6. Limitations of the present meta-analysis

Specific characteristics of meta-analyses in general should be kept in mind when discussing the results. Once again we would like to point out that the present meta-analysis is limited to studies including animations which are not interactive or just minimally interactive, which is a necessary prerequisite for fair comparisons. In addition, the included studies had to fulfill further requirements: they had to compare animated displays with static displays while not mixing both types of visualization, they had to offer two at least roughly informationally equivalent versions, and they had to specify the basic statistics needed for computing effect sizes. As described above, these criteria were the reasons for excluding a number of studies, which might have had potentially interesting results.

Finally, the present meta-analysis does not claim to have included all potentially relevant moderator variables. Quite to the contrary, there are many possible other variables that could be considered, such as prior knowledge (e.g., ChanLin, 2001), spatial ability (e.g., Yang et al., 2003), motivation (e.g., Höffler, 2003), the number of displayed key pictures (e.g., Hegarty, 1992), learners' time on task (e.g., Tversky et al., 2002) or, of course, the option that the learner may interact with the animation (e.g., Nerdel, 2003; Ploetzner & Lowe, 2004; Schnotz, Bockheler, & Grzondziel, 1999; Schuh, Gerjets, & Scheiter, 2005). Because these and other potential factors of instructional relevance have been analyzed or mentioned only by a small number of studies, they could not be included as moderator variables in the present meta-analysis.

6. Conclusion

The present meta-analysis is an attempt to bring some objectivity into a field of research that is still quite difficult to survey. With a mean weighted effect size of $d = 0.37$ and a 95% confidence interval, 95% CI, from 0.25 to 0.49, we found a rather substantial overall advantage of animations over static pictures, and this advantage becomes particularly evident under specific combinations of instructionally relevant circumstances.

There is evidence that animations are specifically superior to static pictures when the depicted motion in the animation explicitly refers to the topic to be learned (i.e. when the visualization plays a representational role; $d = 0.40$, 95% CI 0.26–0.53). However, when the visualization is intended to play a decorational rather than a representational role, animations are not superior to static pictures ($d = -0.05$, 95% CI −0.37 to 0.27). Furthermore, evidence was found suggesting the use of an adequate level of realism (e.g., video-based animations; $d = 0.76$, 95% CI 0.39–1.13). Finally, there is evidence that animations seem to be especially effective for acquiring procedural-motor knowledge ($d = 1.06$, 95% CI 0.72–1.40), but as well for acquiring declarative ($d = 0.44$, 95% CI 0.27–0.57) or problem-solving knowledge ($d = 0.24$, 95% CI 0.04–0.44). Note, however, that the effects of representational role, realism and psycho-motor type of knowledge cannot satisfactorily be decomposed because these features of animations are often combined and thus do not vary independently across the primary studies of the present meta-analysis.

It is desirable to validate the results of the present meta-analysis using additional primary studies that may be published in the future. However, the results of our analysis suggest that animations are capable – in specific areas, under specific circumstances – of facilitating learning even without being interactive. Thus, animations appear to be better than their reputation, but – in order to be effective – they require an instructional design that is grounded in research-based theories about learning and instruction. In other words, we agree with Bétrancourt and Tversky (2000) that an animation is "not a panacea in itself" (p. 326).

References

Ainsworth, S., & VanLabeke, N. (2004). Multiple forms of dynamic representation. *Learning and Instruction, 14*, 241–255.

Alesandrini, K. L. (1982). Imagery-eliciting strategies and meaningful learning. *Journal of Mental Imagery, 6*, 125–140.

Baddeley, A. (1992). Working memory. *Science, 255*, 556–559.

* Baek, Y. K., & Layne, B. H. (1988). Color, graphics, and animation in a computer assisted learning tutorial lesson. *Journal of Computer-Based Instruction, 15*, 131–135.

Bangert-Drowns, R. L. (1986). Review of developments in meta-analytic methods. *Psychological Bulletin, 99*, 388–399.

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: a meta-analysis. *Review of Educational Research, 74*, 29–58.

Bétrancourt, M., & Tversky, B. (2000). Effect of computer animation on users' performance: a review. *Travail-Humain, 63*, 311–329.

* Blake, T. (1977). Motion in instructional media: some subject-display mode interactions. *Perceptual and Motor Skills, 44*, 975–985.

Carlson, R., Chandler, P., & Sweller, J. (2003). Learning and understanding science instructional material. *Journal of Educational Psychology, 95*, 629–640.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review, 14*, 5–26.

* Catrambone, R., & Seay, A. F. (2002). Using animation to help students learn computer algorithms. *Human Factors, 44*, 495–511.

Chandler, P. (2004). The crucial role of cognitive processes in the design of dynamic visualizations. *Learning and Instruction, 14*, 353–357.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293–332.

* ChanLin, L.-J. (1998). Animation to teach students of different knowledge levels. *Journal of Instructional Psychology, 25*, 166–175.

ChanLin, L.-J. (2000). Attributes of animation for learning scientific knowledge. *Journal of Instructional Psychology, 27*, 228–238.

* ChanLin, L.-J. (2001). Formats and prior knowledge on learning in a computer-based lesson. *Journal of Computer-Assisted Learning, 17*, 409–419.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*, 149–210.

Cohen, J. (1966). Some statistical issues in psychological research. In B. B. Wolman (Ed.), *Handbook of clinical psychology* (pp. 95–121). New York: McGraw Hill.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Lawrence Erlbaum Associates.

* Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology, 94*, 428–434.

Dwyer, F. M. (1978). *Strategies for improving visual learning*. Pennsylvania: Learning Services.

Erez, A., Bloom, M. C., & Wells, M. T. (1996). Using random rather than fixed effect models in meta-analysis: implications for situational specificity and validity generalization. *Personnel Psychology, 49*, 275–306.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction, 15*, 313–331.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills: Sage.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: a theory of cognitive interest in science learning. *Journal of Educational Psychology, 90*, 414–434.

* Hays, T. A. (1996). Spatial abilities and the effects of computer animation on short-term and long-term comprehension. *Journal of Educational Computing Research, 14*, 139–155.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando: Academic Press.

* References marked with an asterisk indicate studies included in the meta-analysis.

Hegarty, M. (1992). Mental animation: inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 18*, 1084–1102.

Hegarty, M. (2004). Dynamic visualizations and learning: getting to the difficult questions. *Learning and Instruction, 14*, 343–351.

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.

*Höffler, T. N. (2003). Animation, Simulation oder Standbilder? Zur Wirkung interaktiver Animationen auf das Verstaendnis komplexer biologischer Prozesse [Animation, simulation or static pictures? Concerning the effectiveness of interactive animations upon the comprehension of complex biological processes]. Unpublished diploma thesis, University of Kiel, Germany.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*, 65–70.

Hunter, J. E., & Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis methods: implications for cumulative research knowledge. *International Journal of Selection and Assessment, 8*, 275–292.

Jaccard, J. (2006). *Zumastat 4.0* [Computer software]. Miami: Applied Scientific Analysis.

* Kaiser, M. K., Proffitt, D. R., & Anderson, K. (1985). Judgments of natural and anomalous trajectories in the presence and absence of motion. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 11*, 795–803.

* Lai, S.-L. (2000). Increasing associative learning of abstract concepts through audiovisual redundancy. *Journal of Educational Computing Research, 23*, 275–289.

Large, A. (1996). Computer animation in an instructional environment. *Library & Information Science Research, 18*, 3–23.

Large, A., Beheshti, J., Breuleux, A., & Renaud, A. (1996). Effect of animation in enhancing descriptive and procedural texts in a multimedia learning environment. *Journal of the American Society for Information Science, 47*, 437–448.

Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: a review of research. *Educational Communication and Technology Journal, 30*, 195–232.

Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In D. M. Willows, & H. A. Houghton (Eds.), *The psychology of illustration: I. Basic research* (pp. 51–85). New York: Springer.

Lewalter, D. (1997). *Lernen mit Bildern und Animationen: Studie zum Einfluss von Lernermerkmalen auf die Effektivitaet von Illustrationen* [*Learning with pictures and animations: a study on the effects of student attributes on the effectiveness of illustrations*]. Muenster: Waxmann.

* Lewalter, D. (2003). Cognitive strategies for learning from static and dynamic visuals. *Learning and Instruction, 13*, 177–189.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education, 14*, 225–244.

Lowe, R. K. (2003). Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction, 13*, 157–176.

* McCloskey, M., & Kohl, D. (1983). Naive physics: the curvilinear impetus principle and its role in interactions with moving objects. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 9*, 146–156.

Mayer, R. E. (2001). *Multimedia learning*. Cambridge: Cambridge University Press.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge: Cambridge University Press.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*, 715–726.

Mayer, R. E., & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational Psychology Review, 14*, 87–99.

* Michas, I. C., & Berry, D. C. (2000). Learning a procedural task: effectiveness of multimedia presentations. *Applied Cognitive Psychology, 14*, 555–575.

Milheim, W. D. (1993). How to use animation in computer-assisted-learning. *British Journal of Educational Technology, 24*, 171–178.

*Nerdel, C. (2003). Die Wirkung von Animation und Simulation auf das Verstaendnis von stoffwechselphysiologischen Prozessen [The influence of animation and simulation on the comprehension of physiological metabolism]. Doctoral dissertation, University of Kiel, Germany.

* Nicholls, C., & Merkel, S. (1996). The effect of computer animation on students' understanding of microbiology. *Journal of Research on Computing in Education, 28*, 359–371.

Orwin, R. G. (1983). A fail-safe *N* for effect size in meta-analysis. *Journal of Educational Statistics, 8*, 157–159.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.

Park, O.-C., & Hopkins, R. (1993). Instructional conditions for using dynamic visual displays – a review. *Instructional Science, 21*, 427–449.

Ploetzner, R., & Lowe, R. (2004). Editorial: dynamic visualization and learning. *Learning and Instruction, 14*, 235–240.

Quintana, S. M., & Minami, T. (2006). Guidelines for meta-analyses of counseling psychology research. *The Counseling Psychologist, 34*, 839–877.

* Rieber, L. P. (1989). The effects of computer animated elaboration strategies and practice on factual and application learning in an elementary science lesson. *Journal of Educational Computing Research, 5*, 431–444.

* Rieber, L. P. (1990). Using computer animated graphics with science instruction with children. *Journal of Educational Psychology, 82*, 135–140.

* Rieber, L. P. (1991). Animation, incidental learning, and continuing motivation. *Journal of Educational Psychology, 83*, 318–328.

Rieber, L. P. (1994). *Computers, graphics, and learning*. Madison: Brown & Benchmark.

* Rieber, L. P., Boyce, M. J., & Assad, C. (1990). The effects of computer animation on adult learning and retrieval tasks. *Journal of Computer-Based Instruction, 17*, 46–52.

Rieber, L. P., & Kini, A. S. (1991). Theoretical foundations of instructional applications of computer-generated animated visuals. *Journal of Computer-Based Instruction, 18*, 83–88.

* Rigney, J. W., & Lutz, K. A. (1976). Effect of graphic analogies of concepts in chemistry on learning and attitude. *Journal of Educational Psychology, 68*, 305–311.

Rosenthal, R. (1979). The "file-drawer problem" and tolerance for null results. *Psychological Bulletin, 86*, 638–641.
Rosenthal, R. (1991). *Meta-analytic procedures for social research.* Newbury Park: Sage.
Salomon, G. (1979). *Interaction of media, cognition, and learning: An exploration of how symbolic forms cultivate mental skills and affect knowledge acquisition.* San Francisco: Jossey-Bass.
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., Bockheler, J., & Grzondziel, H. (1999). Individual and co-operative learning with interactive animated pictures. *European Journal of Psychology of Education, 14*, 245–265.
Schuh, J., Gerjets, P., & Scheiter, K. (2005). Fostering the acquisition of transferable problem-solving knowledge with an interactive comparison tool and dynamic visualizations of solution procedures. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th annual conference of the cognitive science society* (pp. 1973–1978). Mahwah, NJ: Lawrence Erlbaum Associates.
* Spangenberg, R. W. (1973). The motion variable in procedural learning. *AV Communication Review, 21*, 419–436.
* Spotts, J., & Dwyer, F. M. (1996). The effect of computer-generated animation on student achievement of different types of educational objectives. *International Journal of Instructional Media, 23*, 365–375.
Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295–312.
* Swezey, R. W. (1991). Effects of instructional strategy and motion presentation conditions on the acquisition and transfer of electromechanical troubleshooting skill. *Human Factors, 33*, 309–323.
* Szabo, M., & Poohkay, B. (1996). An experimental study of animation, mathematics achievement, and attitude toward computer-assisted instruction. *Journal of Research on Computing in Education, 28*, 390–402.
Tallmadge, G. A. (1977). *The joint dissemination review panel ideabook*. Washington, DC: National Institute of Education & U.S. Office of Education.
Tversky, B., Morrison, J.-B., & Bétrancourt, M. (2002). Animation: can it facilitate? *International Journal of Human Computer Studies, 57*, 247–262.
Weidenmann, B. (2002). Abbilder in Multimediaanwendungen [Pictures in multimedia applications]. In L. J. Issing, & P. Klimsa (Eds.), *Information und Lernen mit Multimedia und Internet* (pp. 83–96). Weinheim: Beltz-PVU.
Weiss, R.-E., Knowlton, D.-S., & Morrison, G.-R. (2002). Principles for using animation in computer-based instruction: theoretical heuristics for effective design. *Computers in Human Behavior, 18*, 465–477.
Wittrock, M. C. (1974). Learning as a generative process. *Educational Psychologist, 11*, 87–95.
Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist, 24*, 345–376.
* Wright, P., Milroy, R., & Lickorish, A. (1999). Static and animated graphics in learning from interactive texts. *European Journal of Psychology of Education, 14*, 203–224.
* Yang, E. M., Andre, T., & Greenbowe, T. Y. (2003). Spatial ability and the impact of visualization/animation on learning electrochemistry. *International Journal of Science Education, 25*, 329–349.