When and why does animation enhance learning? A meta-analysis

Sandra Berney, University of Geneva, TECFA, Switzerland

Mireille Bétrancourt, University of Geneva, TECFA, Switzerland

Does the use of animation in a learning context improve comprehension and / or knowledge acquisition of dynamic phenomena, compared with static illustrations? Previous research on this question led to inconsistent findings, depending on various factors related to the learning content, the learning situation, or the delivery features. Hence, a better question to ask is: when and why do animations improve learning? This entails identifying the conditions as well as the associated underlying cognitive processes for which animated graphics may be beneficial.

1. Goal of the study

This study examined the status of the use of multimedia animation in learning environments, by taking into account the cognitive processes underlying its efficiency. Previous literature reviews (Bétrancourt and Tversky, 2000; Hegarty, Kriz & Cate, 2003, Schneider, 2007), which adopted a descriptive or qualitative analysis, report **inconclusive findings** regarding the benefit of animation compared to static graphics. The authors pointed out several problems which are most often encountered. Among them, the unequal amount of information conveyed by both display (Bétrancourt & Tversky, 2000). The level of interactivity and particularly the possibility to control the pace of the animated sequence (e.g., Fischer, Lowe & Schwan, 2006), as well as the task's tests used for measuring the learning comprehension also varied across studies. Schneider (2007) underlines that within the same study, the effect of animation could differ depending on the learning outcomes tested. For example, in a time zones study (Schnotz, Böckheler and Grzondziel, 1999) the comprehension test determined an advantage of the animation over static graphics, whereas the mental simulation test showed the opposite pattern. However, Schneider's review (2007) used a qualitative approach that does not allow the precise identification of factors mediating the effectiveness of animation. The author used a factorial analysis and found varying effects depending on domains, but did not suggest any explanation on the reasons why.

A quantitative approach, through a meta-analysis procedure, will bring a different view of multiple studies, and help clarify and define their strengths and weaknesses. The meta-analysis procedure allows synthesizing a large number of pair-wise comparisons. Its advantage over a qualitative review is that it standardizes findings across studies such that they can be directly compared (Wilson, 2003). It is based on a statistical combination of the results (outcomes) of several studies selected on the basis of precise inclusion/exclusion criteria, such as the definition of the study objective, its population, the group modalities, the outcome measures, and some other literature domain relevant features. The dependant variables (outcomes) of the studies are represented by a normalized measure: the effect-size (ES), which is the result of the standardized mean difference between the 2 compared groups. The computation of the effect-sizes provides an overall finding effect. Crucial to any

meta-analysis is the assumption of homogeneity among the collected data: are the effects independent from one another? However, the presence of heterogeneity can have several explanations. It can be statistical (non concordance of studies outcomes) and/or could be explained by common moderating characteristics underlying the dependant variable. In addition to the global effect-size calculation, the heterogeneity allows us to carry out sub-groups analyses, which would make it possible to highlight nonsignificant effects and therefore indicates the moderator variables influencing the global effect (Cucherat, 2002). While this method is hardly ever used in the multimedia instructional domain, it is frequently used in medicine, for example. To the best of our knowledge, only three meta-analyses were conducted: on the modality effect (Ginns, 2005), on the spatial and temporal contiguity (Ginns, 2006) and the last one, on the animated versus static display instructional learning (Hoeffler & Leutner, 2007). Using this method, Hoeffler & Leutner (2007) found an overall beneficial effect of animation over static graphics.

In this paper, we propose to summarize the large and complex current literature on the use and/or effectiveness of animations in a learning environment. Our central focus of interest is whether instructional materials containing animations are beneficial to the learning and/or comprehension of dynamic phenomena compared to static graphic displays. The comparison of the outcomes between the animated and static graphics groups will provide estimates of the advantage of one particular display, and may moreover explain that the global effect depends upon specific characteristics.

2. What is animation?

Many different definitions of the term animation coexist. Roncarrelli (1989) defined the animation as a “series of single frames, each showing incremental changes in the position of the subject image which when shown in sequence, at high speed, give the illusion of movement. Scaife & Rogers (1996) characterized the movement as spatial and temporal changes. Static images (or frames) and the movement, given by the speed, are the two common point of these authors. A completing suggestion was expressed by Bétrancourt and Tversky (2000) according to which the user can interact with the display. These authors defined animation as “any application which generates a series of frames, so that each frame appears as an alteration of the previous one, and where the sequence of frames is determined either by the designer or the user”(p. 313).

An animation may provide more information for the learner, since it offers a direct visualization of the microsteps that are the minute changes occurring in a dynamic system, which have to be inferred from a series of static graphics (Bétrancourt, Morisson and Tversky, 2001). Another clear advantage is its ability to convey changes over time (Schnotz & Lowe, 2003). Conversely, a series of simultaneously presented static graphics allows for the different states or steps within a depicted process to be consulted at any time, while an animation must be repeated as a whole (Bétrancourt et al., 2001).

Lowe (1999) distinguishes three dynamic changes depicted in an animation: translation, transformation and transition, which focus on position, form and inclusion changes respectively. Schnotz (2005) identified several functions of animations which influence positively the learning. The enabling effect (Schnotz & Rash, 2005) translates the beneficial potential of an animation which allows the learner to process more information than with static graphics. This effect is akin to the supplantation effect (Salomon, 1994), which is an

external cognitive aid for mental operations or processes. The facilitating function (Lowe & Schnotz, 2008) helps the learner to build a dynamic mental representation of the analogous dynamics of the phenomenon, but may lead to an “illusion of understanding”.

Instructional effects of animation may not always be beneficial (Schnotz, Böckheler &Grzondziel, 1999). The animations must be adequately processed in order to extract the necessary information for a coherent mental model building. The subprocesses - selection, organisation, and integration (Mayer, 2005) - may be hindered because the multiple components of the dynamic structure compete with the learner’s attention (Lowe, 1999). Another difficulty is the transience of the animation, which by definition, has to be processed in motion where it could be difficult to perceive all the simultaneous basic changes (Bétrancourt et al., 2001).

Several multimedia learning theories explain the information processes from an animation. The cognitive load theory (Chandler & Sweller, 1991, Sweller, 1988, Sweller & Chandler, 1994, Sweller, van Merrienboer & Paas, 1998) highlights the implications of multimodal information processing on the cognitive structures. Multimedia learning theory from Mayer and Moreno (Mayer, 2001; Moreno & Mayer, 1999) emphasizes the efficiency of learning that is based on a simultaneous presentation of verbal and visual materials. Schnotz and Bannert’s integrated model of text and picture comprehension (Schnotz, Böckheler & Grzondziel, 1999, Schnotz & Bannert, 2003) allows to understand the processes involved in the comprehension of bimodal documents, and explains how integration and collaboration processes of visual and verbal information are necessary (compulsory) in new knowledge building. The description of processes involved in the dynamic phenomenon comprehension building is illustrated by the multimodal comprehension model of Narayanan and Hegarty’s (1998).

3. A meta-analysis to highlight the features that influence the animation process (moderator factor power)

According to the domain literature, several animation characteristics may influence the comprehension and/or the learning of dynamic phenomena.

- The level of dynamic changes within the animation (Lowe, 2003)
- The potential control over the animation that is the pacing, and the interactivity of the display which modifies the animation depending in the user’s activity (Mayer & Chandler, 2001).
- The learner’s cognitive abilities determine the building of a coherent mental model from animation (Lowe, 2003)
- The impact of prior knowledge level on learning with animations varies across learners (ChanLin, 1998, 200, Kalyuga, 2008).
- The learners’ spatial abilities level influence the animation efficiency processing (Hegarty & Sims, 1994, Yang et al., 2003).
- The type of knowledge and instructional domain (Bétrancourt et al., 2000)

In line with the issues discussed above, this study aimed to investigate through *a meta-analysis* whether animations as instructional displays are efficient in a learning context and which underlying cognitive processes might be associated with it.

4. Method

Literature search

Starting with the studies selected in Hoeffler & Leutner's study (2007), the literature search was expanded and updated by a systematic search for studies published up to October 2008 comparing animated versus static graphic displays of dynamic phenomena. This was done through the PsycInfo (1806 – 2008), ERIC (1966 – 2007), Francis (1984 – 2007), MedLine (1950 – 2008) and Psyndex (1945 – 2008) databases. Following keywords were searched; *animation, multimedia, interactive animation, static graphic, multimedia learning, dynamic picture, static picture.* Reviews studies mentioned earlier in this paper were a great help in order not to miss any essential study.

Eligibility criteria

Studies were primarily selected based on their abstract. Secondly, they were only included in our review when the following conditions were fulfilled : (a) the empirical study evaluated the impact of different instructional format displays, in particular a comparison of animated versus static graphic display, (b) the studies must have been written in English, French or German, (c) the dependant variables had to be an operationalization of the knowledge comprehension of the learners and (d) the study had to provide sufficient descriptive data to calculate the effect size (if *means, standard deviations, or F* were not mentioned, we tried to contact the author to get the missing data). Based on these criteria, 53 articles (63 studies) were selected for this review. 3 articles were excluded because we did not get the missing information. It reduced the number of articles to 50 and the number of experiments to 59.

Variables coded from the studies

For each available experiment, the features were coded according to three different type of information: 1. study identity, 2. moderator factors and 3. learning comprehension outcomes.

Nine following variables were extracted:

a) *authors and year of publication*
b) *sample size*
c) *sample age (*pupils, undergrad students, graduate students, adults)
d) *learning instruction domain specificity* (aeronautics, astronomy, biology, chemistry, cognitive events, geography, geology, informatics, mathematics, mechanics, meteorology, physics)
e) *prior knowledge measure*
f) *spatial abilities measure*
g) *presence of display control in the animation.*
h) *dependant variables data* (For this analysis, only tests used for knowledge comprehension assessment were examined) and *their statistical values*
i) *principal outcomes of the dependant variables*

Methodological problems arose while extracting the dependant variables. In fact, the construct of interest, which is the knowledge comprehension assessment, was measured in the studies with many different tests.

This outcomes' variety brought up an important question: do the various comprehension tests measure the learner's knowledge comprehension? In addition, the multiple outcomes measures within studies violated the assumption of independence, central for the meta-analysis approach.

We faced these problems by applying a correction on the sample size, and also, by recoding each outcome with a standard criterion based on the revised version of Bloom's taxonomy (Anderson and Krathwohl, 2001). This last correction added one more variable *j)*, for a total of ten.

j) *knowledge outcome according to the Bloom revised taxonomy*

Recoding of the multiple dependant variables

In our research, on 50 studies, 14 different types of learning tests were mentioned, such as *retention, inference, near & far transfer, multiple choice questionnaire, comprehension, problem solving, fill in the blanks, drawing, descriptive learning, free recall, procedural, preferences and matching tests*. Each researcher or author built his own test to analyze the dependant variable. No harmony in the measurement of the learners' comprehension of the dynamic phenomenon in a multimedia learning context seems to emerge, as could be seen in other psychology domains. Although the validity and the fidelity of the "home-made" tests are not questioned, it is difficult to compare the outcomes of these studies, particularly the comprehension level and the learning depth of the learners.

Very few studies (4 out of 50) mentioned the correspondence between their dependant variable outcomes and a knowledge standard criterion. Rieber (1898, 1990, 1991a) mentioned Gagné's taxonomy (Gagné, 1985) and Rigney (1976) the Bloom one (Bloom, 1956). The use of a pedagogical model as standard knowledge acquisition criteria may be interesting due to its classification of knowledge acquisition levels. Many pedagogical models and taxonomies are defined in the literature. The Bloom revised taxonomy (Anderson and Krathwohl, 2001) was chosen because it highlights the link between knowledge and cognitive abilities. Four levels are described in the *knowledge dimension*, which are *factual*, *conceptual*, *procedural* and *meta-cognitive.* Characterized by their complexity, six levels of the cognitive abilities dimension are defined, which are *remember, understand, apply, analyze, evaluate* and *create*. This revised taxonomy may be seen as a double-entry graphic, where the cognitive dimension crosses the knowledge dimension. The intersection of each entry translates the cognitive process involved in the learning process. With the use of the Bloom revised taxonomy (table 1), we could decrease the quantity of different outcomes, from 13 to 9.

Table 1 : Number of pair-wise comparisons of the sample falling into each category of the learning outcomes as defined in the revised Bloom's taxonomy (Anderson and Krathwohl, 2001).

	Cognitive processes dimension				
Knowledge level					
	remember	**understand**	**apply**	**analyze**	
factual	20	23			
conceptual	20	49	28	4	
procedural	1	9	17		
	41	81	45	4	171

Calculation

Each outcome was reported as an independent statistical value (Table 2). The effect sizes measures for two independent groups were calculated from means and standard deviation directly reported in the studies, except in 21 comparisons that were based on *F* statistics and on Chi statistics for 2 of them (formula to compute effect sizes from other statistics than means and SD were adapted according to Lipsey & Wilson (2001) procedures.) The effect sizes were computed as Hedges's g (standardized mean difference effect size) based on standardized difference, which defines a variation on Cohen's *d*, (Lipsey & Wilson, 2001). Hedges's formula of effect size is $ES_{sm} = \frac{M_{anim} - M_{static}}{S_{pooled}}$. S_{pooled} is defined as the square root of the average of the squared standard deviations. A bias correction for small sample sizes was adopted as: $g = \left[1 - \frac{3}{4N-9}\right] ES_{sm}$. To avoid an over representation of large sample size effect sizes, g was weighted by the inverse of standard error (Lipsey & Wilson, 2001). To correct the case of multiple dependant variables within studies, we followed the correction described in Hoeffler and Leutner's study (2007): within a study with multiple outcomes, the ES of each outcome was weighted by adjusting the sample size according to the number of non-independent pair-wise comparisons of this study. The outliers were detected according to Huffcutt & Arthur's procedure (cited by Lipsey & Wilson, 2001), which determined them as a break in the effect size distribution. Consequently, 2 effect sizes in Lowe's study (2003) were found to have a higher value than -4. They were brought back to the less extreme value (-3.88). From a conventional usage, a positive value of g demonstrates a beneficial efficiency of the animations.

The variability of the overall effect sizes was testes with an homogeneity analysis, a Q statistics which has a χ^2 (Chi square) distribution (Cooper, 1989, Hedges & Olkin, 1985). To take into consideration the non-homogeneity of our main overall ES, a random-effect model analysis on the multiple dependant variables as well as on the moderators' sub-groups was adopted. The random-effect model assumes that each study is associated with a different but related parameter (Normand, 1999). The procedure of that model is that each comparison is weighted by the inverse of the sampling variance plus a constant that represents the variability across the population effects (Wilson, 2006). A weighted regression analysis

(Lipsey & Wilson, 2001) was calculated to assess the relationship between ES and the different moderators. It explains the excess variability in the effect size distribution.

The calculations of the standardized mean difference effect sizes, the random-effect model and the weighted regression analyses were realized with two macros (Wilson, 2005) ran on the SPSS © software. The MEANS macro computed meta-analytic summary statistics, and METAR computed a weighted regression procedure.

5. Results

5.1. Overall analysis

One hundred and seventy-one effects derived from between-group experiments were extracted from these studies (cf. table 1). The total number of participants across the studies was 8375.

```
Run MATRIX procedure:

Version 2005.05.23

*****  Meta-Analytic Results  *****

------- Distribution Description --------------------------------
           N      Min ES      Max ES    Wghtd SD
     171.000      -3.880       4.593        .809

------- Fixed & Random Effects Model ----------------------------
         Mean ES    -95%CI    +95%CI        SE         Z         P
Fixed      .3197     .2549     .3845     .0331    9.6675     .0000
Random     .2899     .1519     .4279     .0704    4.1166     .0000

------- Random Effects Variance Component ------------------------
v    =    .474641

------- Homogeneity Analysis -------------------------------------
           Q          df           p
    598.3057    170.0000       .0000

Random effects v estimated via noniterative method of moments.

------ END MATRIX -----
```

Figure 2 : Matrix of the mean standardized ES after the outliers correction.

The overall standardized mean effect size from the random effects model showed a significant value 0.2899 (p< 0.0001, 95 % confidence interval 0.15 – 0.42). According to Cohen's rule of thumb (1988), the overall ES has a small magnitude. This result suggests that the use of animation in learning dynamic phenomena is beneficial compared to static graphic display. The test on homogeneity shows that the null hypothesis can be rejected (Q = 598.3, df=170, p< 0.01). It indicates that one of more moderator characteristics, other than sampling error or chance, may account for this heterogeneity.

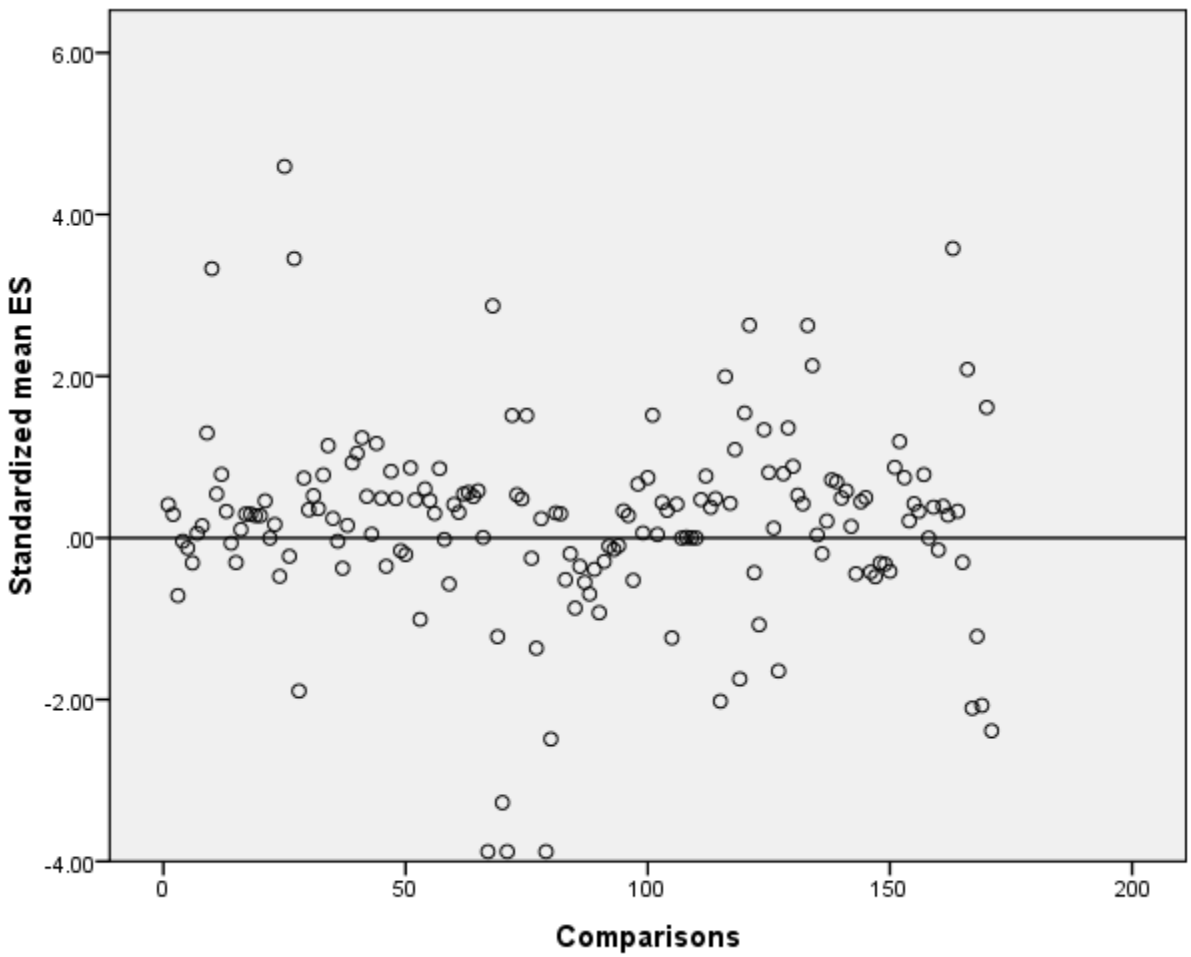

Figure 3 : Scatter diagram of the ESs

Of the 171 pair-wise comparisons in this synthesis (cf. figure 3), 110 (64, 3%) of the ES were positive and favored the animated display group, 60 (35,1 %) were negative and favored the static display group and 1 (0,6 %) was null. More than the half of the ESs (53,8%) lies between the values -0.5 and 0.5.

5.2. Multiple dependant variables analysis

Effect sizes for the multiple dependant variables are given in Table 2.

Table 2 : summary of the outcomes' ES according to a random-effect model analysis. The mean of the standardized mean difference, the 95 % interval confidence, *Z, p* and Q are shown.

Bloom Taxonomy revised conditions	**Mean ES**	**95 % Confidence Interval**	**Z**	***p***	**Q**	***n***
Multiple Dependant variables						
Factual – remember	0.5849 *	0.3569 – 0.8129	5.0273	0.0000	13.094	20
Factual – understand	-0.0352	-0.2256 – 0.1552	-0.3621	0.7173	16.804	23

Conceptual – remember	0.5076	-0.0866 – 1.1019	1.6744	0.0941	114.015	20
Conceptual – understand	0.2879 *	0.0292 – 0.5466	2.1812	0.0292	215.715	49
Conceptual – apply	0.1305	-0.0148 – 0.2758	1.7602	0.0784	28.8188	28
Conceptual – analyze	-0.0054	-0.6268 – 0.6159	-0.0172	0.9863	1.641	4
Procedural –understand	0.4488	-0.3673 – 1.2650	1.0788	0.2811	71.914	9
Procedural – apply	0.5406 *	0.0386 – 1.0427	2.1106	0.0348	21.5936	17
Knowledge dimension						
Factual	0.2184 *	0.0577 – 0.3790	2.6632	0.0077	46.6362	43
Conceptual	0.2686 *	0.0987 – 0.4385	3.0981	0.0019	276.584	101
Procedural	0.4572	-0.1275 – 1.0418	1.5325	0.1254	162.229	27
Cognitive processes dimension						
Remember	0.6058 *	0.2401 – 0.9715	3.2469	0.0012	182.054	41
Understand	0.2168 *	0.0220 – 0.4116	2.1816	0.0291	313.971	81
Apply	0.1967 *	0.0244 – 0.3691	2.2369	0.0253	64.2630	45
Analyze	-0.0054	-0.6268 – 0.6159	-.0172	0.9863	1.641	4

* significant $p < 0.05$;

Pair-wise comparisons between the types of outcomes (the different categories of the revised taxonomy of Bloom (table 1)) indicated differences between categories of the dependant variable (cf. table 2). Significant positive ESs for the factual – remember (g = 0.5849, 95% CI 0.35 – 0.81), for conceptual – understand (g = 0.2879, 95% CI 0.29 – 0.54) and for procedural – apply categories (g = 0.5406, 95% CI 0.03 – 1.04) were found.

Pair-wise comparisons between the knowledge dimension indicated that both the factual (g = 0.2184, 95% CI 0.05 – 0.37) and the conceptual (g = 0.2686, 95% CI 0.09 – 0.43) dimensions were significant. Among the cognitive processes dimension, the skills remember (g = 0.6058, 95% CI 0.24 – 0.97) and understand (g = 0.1967, 95% CI 0.02 – 0.36) were significant.

5.3 Moderators

As mentioned in the introductory part of this paper, the choice of moderator characteristics was made according to the multimedia learning models and the current literature.

Concerning the moderator factors (cf. table 3), the standardized mean ES was significant and greater for high spatial abilities (g = 0.6919, 95% CI 0.16 – 1.24) than low spatial abilities (g = 0.2314, 95% CI 0.09 – 0.36). The inverse pattern existed for the control display: groups with no display had a greater ES (g = 0.3445, 95% CI 0.15 – 0.53) than groups with control (g = 0.2249, 95% CI 0.04 – 0.40). For prior knowledge, only the novice group (g = 0.2573, 95% CI 0.15 – 0.35) showed a significant ES. If we compared the sample range, the pupils showed a greater ES (g = 0.6146, 95% CI 0.12 – 1.10) than the whole adult groups

(g = 0.2386, 95% CI 0.10 – 0.36). For the instructional knowledge specificity, three out of twelve domains showed a significant result. ESs in descending order are Astronomy (g = 0.4309, 95% CI 0.11 – 0.74), Mathematics (g = 0.3876, 95% CI 0.00 – 0.77) and Informatics (g = 0.2719, 95% CI 0.07 – 0.46).

Table 3: summary of the moderators' ES according to a random-effect model analysis. The mean of the standardized mean difference, the 95 % interval confidence, *Z, p* and Q are shown.

Moderators	**Mean ES**	**95 % Confidence Interval**	**Z**	***p***	**Q**	***n***
Spatial abilities						
High	0.6919 *	0.1680 – 1.2451	2.5889	90.0096	149.827	21
Low	0.2314 *	0.0956 – 0.3672	3.3406	0.0008	431.283	150
Prior knowledge						
High	0.6499	-0.2738 – 1.5737	1.3790	0.1679	341.356	14
Low	0.2573 *	0.1583 – 0.3563	5.0933	0.000	244.600	157
Control display						
Yes	0.2249 *	0.0465 – 0.4034	2.4706	0.0135	108.904	65
No	0.3445 *	0.1589 – 0.5301	3.6372	0.0003	488.832	106
Sample range						
Pupils	0.6146 *	0.1246 – 1.1045	2.4586	0.0139	241.513	35
Undergraduates	0.2664 *	0.0850 – 0.4478	2.8778	0.004	94.21	59
Students	0.1977 *	0.0159 – 0.3794	2.1316	0.0330	247.275	74
Adults	0.7661 *	0.3263 – 1.2058	3.4144	0.0006	2.815	3
Adults *(undergraduate + students + adults)*	0.2386 *	0.1090 – 0.3681	3.6091	0.0003	351.073	136
Instructional knowledge						
Aeronautics	-0.0042	-0.8428 – 0.8344	-0.0098	0.9922	0.304	4
Astronomy	0.4309 *	0.1173 – 0.7444	2.6935	0.0071	0.107	2
Biology	0.1288	-0.0939 – 0.3515	1.1333	0.2571	31.177	22
Chemistry	1.0861	-0.3640 – 2.5363	1.4680	0.1421	98.960	6
Cognitive events	0.3336	-0.0304 – 0.6977	1.7962	0.0725	46.477	19
Geography	0.2236	-1.0360 – 1.4832	0.3479	0.7279	3.903	2
Geology	-0.6210	-1.2627 – 0.0207	-1.8968	0.0579	0.0505	2
Informatics	0.2719 *	0.0777 – 0.4660	2.7445	0.0061	30.035	20
Mathematics	0.3876 *	0.0036 – 0.7717	1.9784	0.0479	32.192	10
Mechanics	0.1257	-0.0716 – 0.3230	1.2490	0.2117	49.192	32
Meteorology	0.0709	-0.3174 – 0.4591	0.357	0.7206	25.1112	22

Physics	0.6560	-0.0472 – 1.3591	1.8284	0.0675	2221.244	30

* significant p< 0.05 ;

A regression analysis (cf. table 4) was calculated between the three main moderators, such as spatial abilities, prior knowledge and control display.

Table 4: Output from SPSS macro metareg showing mixed model results with 3 moderators – spatial abilities, prior knowledge and control display .

```
Run MATRIX procedure:


 Version 2005.05.23

  *****  Inverse Variance Weighted Regression  *****

  *****  Random Intercept, Fixed Slopes Model  *****

 ------- Descriptives -------
       Mean ES     R-Square            k
         .2900        .0520     171.0000

 ------- Homogeneity Analysis -------
                      Q           df            p
 Model          9.1028       3.0000        .0280
 Residual     165.9471     167.0000        .5085
 Total        175.0499     170.0000        .3793

 ------- Regression Coefficients -------
                  B        SE  -95% CI  +95% CI        Z        P     Beta
 Constant    .2370    .0968    .0473    .4266   2.4487    .0143    .0000
 spatial_    .4915    .2114    .0772    .9057   2.3254    .0200    .1774
 prior_k_    .3312    .2307   -.1209    .7833   1.4357    .1511    .1126
 control_   -.1398    .1546   -.4427    .1632   -.9043    .3658   -.0713

 ------- Method of Moments Random Effects Variance Component -------
 v      =   .47082

 ------ END MATRIX -----
```

The regression model with the three moderators *spatial abilities, prior knowledge* and *control display* explains a part of the variance among moderators (Q= 9.102, *p < .05*). In this model, prior knowledge (Z= 1.435, *p>.05)*, and control display (Z= -0.904, *p>.05)* do not appear to add anything to explaining variability across effect sizes beyond that explained by the spatial abilities (Z= 2.325, *p= .02)*,. The unequal number of studies which measure these moderators may account for this limited result.

DISCUSSION

The results of the meta-analysis support the hypothesis of the beneficial effect of the presence of animated display for learning dynamic phenomena. According to the Bloom's revised taxonomy, learners who were tested on factual or conceptual knowledge had significant better performances as on procedural knowledge, which was not observed in

Hoeffler and Leutner's (2007) meta-analysis. Considering the cognitive dimension, three out of the four processes (remember, understand and apply) were beneficial. The crossed dimensions showed a different pattern, which is difficult to explain. This overall efficiency effect was moderated by the instructional knowledge domains. Students who learned a dynamic phenomenon from astronomy, informatics, or mathematics materials performed better. The learner's spatial abilities level positively correlated with the overall efficiency effect, which suggests that the more animation is found efficient, the more the learner's spatial abilities account for it.

In a more general point of view, these results are difficult to explain, partly because of the methodological problems encountered. There are at least two major inconsistencies that weaken the scientific comparison principle of "all things being equal": the type of learning assessment and the diversity of the animation characteristics, which both vary widely. Theses issues will be further discussed in the presentation.

References

Anderson, L.W. & Krathwohl, D.R. (2001). *A taxonomy for learning, teaching and assessing : a revision of Bloom's taxonomy of educational objectives.* New York, Longman

Benton, A.L., Sivan, A.B., Hamsher, K., Varney, N.R., Spreen, O. (1994). Contributions to neuropsychological assessment. New York: Oxford University Press.

Bétrancourt, M., Bauer-Morrison, J. & Tversky, B. (2001). *Les animations sont-elles vraiment plus efficaces ?* Revue d'intelligence artificielle, 14 *(1-2), 149-166.*

Bloom, B. S. & Krathwohl, D.R. (1956). *Taxonomy of Educational Objectives, the classification of educational goals – Handbook I: Cognitive Domain*. New York: McKay

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

Ekstrom, R.B., French, J.W., Harman, H.H & Dermen, D .(1976) . *Kit of factor-referenced cognitive tests.* Princeton, NJ : Educational Testing Service.

Fischer, S., Lowe, R.K. & Schwan, S. (2007). *Effects of presentation speed of a dynamic visualization on the understanding of a mechanical system.* Applied Cognitive Psychology.

Fischer, S & Schwan, S. (2009). *Comprehending animations: effects of spatial cueing versus scaling.* Learning and Instruction, doi:10.1016/j.learninstruc.2009.05.005

Gagné, E. D. (1985). *The cognitive psychology of school learning.*Boston: Little, Brown.

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

Lipsey, M. W. & Wilson, D.B. (2001). *Practical meta-analysis.* Applied Social Research Methods Series, Vol 49. Thousand Oaks, California : Sage Publications, Inc.

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

Lowe, R, and Schnotz, W. (2008). A unified view of learning from animated and static graphics. In Learning with animation: Research implications for design, 1 Edition, eds Richard Lowe and Wolfgang Schnotz, 304-356. New York: Cambridge University Press.

Mayer, R.E. (2005). Introduction to multimedia learning, In *The Cambridge handbook of multimedia learning*. Cambridge : Cambridge University Press.

Mayer, R.E. (2009). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and Instruction*, doi:10.1016/j.learninstruc.2009.02.012.

Normand, S L. (1999). *Tutorial in biostatics meta-analysis: formulating, evaluating, combining, and reporting.* Statistics in Medicine. *18*, 321-359.

Schnotz, W., Böckheler, J., & Grzondziel, H. (1999). *Individual and co-operative learning with interactive animated pictures*. European Journal of Psychology of Education, *14*, 245–265.

Schnotz, W., & Lowe, R. (2003). *External and internal representations in multimedia learning*, Learning and Instruction, *13*, 117-123.

Wilson, D.B. (2006, June, 13). *David Wilson's statistic home page.* Retrieved June, 10, 2009, from http://mason.gmu.edu/~dwilsonb/

Wilson, D. B. (2005). *Macros for SPSS/Win version 6.1 or higher.* Retrieved June, 10, 2009, from http://mason.gmu.edu/~dwilsonb/ma.html